

580	590	600	610	620	630	640
KNMNVRLPLNF	FEELEFTWLKDKQKNS	FEVGSJDSMDSPYAD	SIKRYISIKSALSAD	KAYENDNMVTLFRSS		
KNMNVRLPLNF	FEELFTWLKDKQKNS	FEVGSJDSMDSPYAD	SIKRYISIKSALSAD	KAYENDNMVTLFRSS		
KNMNVRLPLNF	FEELFTWLKDKQKNS	FEVGSJDSMDSPYAD	SIKRYISIKSALSAD	KAYENDNMVTLFRSS		
580	590	600	610	620	630	640

Initial Score =	751	Optimized Score =	751	Significance =	0.45
Residue Identity =	81%	Matches =	659	Mismatches =	118
Gaps =	0	Conservative Substitutions		<del>28</del>	

[illegible]



```

440      450      460      470      480      490      500
510      520      530      540      550      560      570
HVSNDYSFIRYTRITLYOFQFOEALCOQAKHEGPHKCDISNSTEAGOKLPNMLRLGKSEPPATLALENVYGA
|||||
HVSNDYSFIRYTRITLYOFQFOEALCOQAKHNGSLAKCDISNSTEAGOKLKNLSIGNSEPPTEALEENVYGA
510      520      530      540      550      560      570
580      590      600      610      620      630      640
KNNNVRLNLYFEPLFTWLKDONKNSFVGSTWSPYADQSIKVRISLSKALGDKAYEWNDEMYLFRSSVA
:|||||
RNMVAKPLNTEQPLFDWLKEQNRNSFVGNTNEMSPYADQSIKVRISLSKALGANAYEMTNMEMFLFRSSVA
580      590      600      610      620      630      640
650      660      670      680      690      700      710      720
YAMROYFLAKYKNOMILFGEDEVANLKPRISENFVTAAPKNVSDIIPRTEVEKAIKMSRSKRINDAFRLNDN
|||||
YAMKYSIITKNOTVFEPLLEDVAVSLDKPRVSFYPVTSPONVSDVIPRSEVEDAIRMSRGRINDVFGILNDN
650      660      670      680      690      700      710      720
730      740      750      760      770      780      790
SLEFLGIQPTLGPNOQPVYSIMLIVGVVGVIVGIVLIFTGIDRKKKNNKARSGENPYASIDISKGENN
|||||
SLEFLGIHPTLEPPYQDPVTIMLIIIVGVWALVVGIIILIVTGIGRKKKNETKREENPYDSMDIGKGEEN
730      740      750      760      770      780      790
800      X
PGFQNTDDVQTSF
|||||
AGFQNSDDAQTSF
800      X

```





Pf	29-SEP-1999;	99MO-US22976.
Xx	30-SEP-1998;	98US-0163648.
Pa	(MILL-) MILLENNIUM PHARM INC.	
Pi	Acton LS, Robison KE, Hsieh FY;	
Dx	WPI: 2000-293140/25.	
Dr	N-PSDB; AAA12764.	
Pt	Isolated nucleic acid encoding angiotensin converting enzyme-2 (ACE-2)	
Pt	polypeptide useful for detecting an ACE-2 therapeutic for treating	
Pt	hypertension, congestive heart failure, myocardial infarction,	
Pt	atherosclerosis and renal failure -	
Px	Claim 2; Fig 1; 138pp; English.	
Cc	The present sequence represents a human angiotensin converting enzyme-2	
Cc	(ACE-2). ACE-2 is expressed predominantly in kidneys and testis. The	
Cc	sequence of the full length ACE-2 cDNA was determined from a clone	
Cc	obtained from a cDNA library prepared from mRNA of a human heart of	
Cc	a subject who had congestive heart failure. ACE-2 has significant	
Cc	sequence homologies with ACE enzymes, and has also been shown to	
Cc	hydrolyse angiotensin I into Ang(1-9). The ACE-2 therapeutics are	
Cc	used to treat blood pressure related diseases and conditions, such as	
Cc	hypertension, congestive heart failure, chronic heart failure, acute	
Cc	heart failure, myocardial infarction, atherosclerosis and renal	
Cc	failure.	
Sq	Sequence 805 AA:	
Qy	Query Match 100.0%; Score 4291; DB 21; Length 805;	
Qy	Best Local Similarity 100.0%; Pred. No. 0;	
Matches	805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	1 MSSSSMLLSIAVAAYAAOSTIEQAATFDKRNHEAEDELFYOSLASMYNNTITEENVO 60	
Oy	1 MSSSSMLLSLVAAYTAOSTIEQAATFDKRNHEAEDLFYOSLASMYNNTITEENVO 60	
Oy	61 NNNNGGDKMSAFILKEOSTLAOMYPLOEIONLVKLQLALQONGSSVSEDSKRRLNTLL 120	
Oy	61 NNNNGGDKMSAFILKEOSTLAOMYPLOEIONLVKLQLALQONGSSVSEDSKRRLNTLL 120	
Oy	121 NTMTSTIYGKACNPDFOECLLPBGLMEIMANSIDYNERLAMEWSRSEVGKOLRPXY 180	
Oy	121 NTMTSTIYGKACNPDFOECLLPBGLMEIMANSIDYNERLAMEWSRSEVGKOLRPXY 180	
Oy	181 EEEVLKNDMAANANYIEDYGDYWRGDYEYNGVDYSNGOLIIDEVEHFEELKPLYEHL 240	
Oy	181 EEEVLKNDMAANANYIEDYGDYWRGDYEYNGVDYSNGOLIIDEVEHFEELKPLYEHL 240	
Oy	241 HAYYAKLAMNAVPSYISIPICLPAILLDGMGRFNTNYSLSLVPFGOKPNIDVTAMVDQ 300	
Oy	241 HAYYAKLAMNAVPSYISIPICLPAILLDGMGRFNTNYSLSLVPFGOKPNIDVTAMVDQ 300	
Oy	301 AMDAORIRKEAEKFFVSIGLPNMTOGFWEENSMLDPGNVQRAVCPTAMDLGKGFRTLM 360	
Oy	301 AMDAORIRKEAEKFFVSIGLPNMTOGFWEENSMLDPGNVQRAVCPTAMDLGKGFRTLM 360	
Oy	361 CTKTVMDELFAHHMHGIQYDMAVAAPFLLRNANGSGFHAVGEINSLSATPKNLKS 420	
Oy	361 CTKTVMDELFAHHMHGIQYDMAVAAPFLLRNANGSGFHAVGEINSLSATPKNLKS 420	
Oy	421 IGLSPDFOEDNETEINFNLKQALTIVGTLPFTMLEKRWMMVFEGELPKOOWMKKWEM 480	
Oy	421 IGLSPDFOEDNETEINFNLKQALTIVGTLPFTMLEKRWMMVFEGELPKOOWMKKWEM 480	
Oy	481 KREIVGVPEPVPHDETCDPASLFHVSNDSYFIKYRTTLYOFQFOEALCOAHNEGPLH 540	
Oy	481 KREIVGVPEPVPHDETCDPASLFHVSNDSYFIKYRTTLYOFQFOEALCOAHNEGPLH 540	
Oy	541 KCDISNSGEAQGLENNMLRGKSEPWTLALENVYGAKNMNVRPLINTPEPLFTYLKDQNK 600	

Dd	541	KCDISNSTEAOQKLFNNMLRLGKSEPWTLLENVVGAKKMMNVRPLLNYPEEPLFTMLKQNK	600
Qy	601	NSFVGWSTWDSPPYDQSIKVRISLSKALGDKAYENDNEMTLFSSVAYAMRQYFLKYN	660
Dd	601	NSFVGWSTWDSPPYDQSIKVRISLSKALGDKAYENDNEMTLFSSVAYAMRQYFLKYN	660
Qy	661	QMIFGEBEDVAVANLKPRISENFVYTAAPKNVSDIIPREVEKATIRMSRSRINDAFLRNDN	720
Dd	661	QMIFGEBEDVAVANLKPRISENFVYTAAPKNVSDIIPREVEKATIRMSRSRINDAFLRNDN	720
Qy	721	SLEFLGIQPTLGPNPQPVSIWLVFGVYGVVIVGLIETGIRDRKKKNKARSGENP	780
Dd	721	SLEFLGIQPTLGPNPQPVSIWLVFGVYGVVIVGLIETGIRDRKKKNKARSGENP	780
Qy	781	YASIDISKGENNPGFQNTDYOVSF	805
Dd	781	YASIDISKGENNPGFQNTDYOVSF	805
RESULT 2			
ID	AA67310	standard; Protein; 805 AA.	
XX	AC	AA67310;	
XX	DT	11-APR-2000 (first entry)	
XX	DE	Human MPROT15 amino acid sequence #1.	
XX	KW	MPROT15; treatment; hypertension; human; myocardial disease; apoplexy;	
XX	KW	heart disease; apoplexy; heart disease; nervous denaturation; hormone;	
XX	Alzheimer's	disease; cytokine.	
OS	Homo sapiens.		
XX	PN	JP1318472-A.	
XX	PD	24-NOV-1999.	
XX	PF	22-JAN-1999; 99JP-0014949.	
XX	PR	13-MAY-1998; 98GB-0010373.	
XX	PR	18-AUG-1998; 98GB-0018009.	
XX	PA	(SMIK ) SMITHKLINE BEECHAM PLC.	
XX	DR	WPI: 2000-109268/10.	
XX	DR	N-PSDB; AA259465.	
XX	PT	MPROT15 polypeptide and MPROT15 polynucleotides - useful for the	
XX	PT	treatment of hypertension, myocardial diseases, apoplexy, heart	
XX	PT	diseases, nervous denaturation, Alzheimer's disease etc.	
XX	PS	Claim 1; Page 15; 22pp; Japanese.	
XX	CC	This is amino acid sequence #1 of human MPROT15. The MPROT15	
XX	CC	polynucleotide and polypeptide sequences can be used for the treatment of	
XX	CC	hypertension, myocardial diseases, apoplexy, heart diseases, nervous	
XX	CC	denaturation, Alzheimer's disease and diseases related to the processing	
XX	CC	of peptide hormones and cytokines.	
XX	SQ	Sequence 805 AA;	
Qy	Query Match	100.0%; Score 4291; DB 21; Length 805;	
Qy	Best Local Similarity	100.0%; Pred. No. 0;	
Qy	Matches 805; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Dd	1	MSSSWLLSTVAVAAGOSTIEEAKTFLDKRNHRAEDLFYSSSLASWNYNTNTEENVQ	60
Dd	1	MSSSWLLSTVAVAAGOSTIEEAKTFLDKRNHRAEDLFYSSSLASWNYNTNTEENVQ	60
Qy	61	NNNNAAGDWSAFLKBOSTLAQMYPLQEIQNLTVKQLQALQONGSSVLESDKSRRLNTIL	120

```

Db 61 NMNNAQKMSAFLEKEOSTLAOMYPLQETIQNTLVKLOLQALQONGSSVLSDEKSKRLNTTL 120
OY 121 NTMSTIYSTGKVCNPDNPQECILLEPGLNEIMANSIDYERLMAWESMRSEVGKOLRPLY 180
OY 121 NTMSTIYSTGKVCNPDNPQECILLEPGLNEIMANSIDYERLMAWESMRSEVGKOLRPLY 180
OY 181 EEEVYLKNEKARAHNYEDYDGYWRGDEYVNGVDYDYSRGOLIEDVEHTFEELKPLYEHL 240
OY 181 EEEVYLKNEKARAHNYEDYDGYWRGDEYVNGVDYDYSRGOLIEDVEHTFEELKPLYEHL 240
OY 241 HAYVRAKLMAVAPSYISPIGCLPAHLIDGMWGRFTNLYSLTVPFGOKPNIDVTDAVQ 300
OY 241 HAYVRAKLMAVAPSYISPIGCLPAHLIDGMWGRFTNLYSLTVPFGOKPNIDVTDAVQ 300
OY 301 AMDAORIFKEAEKFEFVSGLPMMTQGFWENSMLTDPGNVOKAVCHPTAMDGKGFRLIM 360
OY 301 AMDAORIFKEAEKFEFVSGLPMMTQGFWENSMLTDPGNVOKAVCHPTAMDGKGFRLIM 360
OY 361 CTKVMTMDEFLTAHHEMGHIQYDMAVAAOPFLLRNGANEHGEHAEIMSLSAATPKHLKS 420
OY 361 CTKVMTMDEFLTAHHEMGHIQYDMAVAAOPFLLRNGANEHGEHAEIMSLSAATPKHLKS 420
OY 421 TGLLSPDQEDNETETINFLKQALTYGTLPTFTMLEKMWVFKGEIPKQOMKKWEM 480
OY 421 TGLLSPDQEDNETETINFLKQALTYGTLPTFTMLEKMWVFKGEIPKQOMKKWEM 480
OY 481 KREIIVGVEPVPHDETYCDPASLFHVSNDYSFIRYTTTLQOPQOPQLQAAKHEGPLH 540
OY 481 KREIIVGVEPVPHDETYCDPASLFHVSNDYSFIRYTTTLQOPQOPQLQAAKHEGPLH 540
OY 541 KCDISNSTAGOKLEMLILGKSEPTLAEVNGAKMMNVRPLNTEPEPLTWLKQONK 600
OY 541 KCDISNSTAGOKLEMLILGKSEPTLAEVNGAKMMNVRPLNTEPEPLTWLKQONK 600
OY 601 NSFVGMSTWSPYADOSIKVRLSLKALCDKAKEMNDENMIFRSVAVAMROYLYLAKYN 660
OY 601 NSFVGMSTWSPYADOSIKVRLSLKALCDKAKEMNDENMIFRSVAVAMROYLYLAKYN 660
OY 661 QMILFGEEDVRANKLPKRSIFNFVAPKRVSDIIPREVEKAKIRMSRINDARFLNDN 720
OY 661 QMILFGEEDVRANKLPKRSIFNFVAPKRVSDIIPREVEKAKIRMSRINDARFLNDN 720
OY 721 SLEFLGIQPLTGPDPNPVSIWLIVFGVWGVIVGIVILITGPIHDKKKKAKSGENP 780
OY 721 SLEFLGIQPLTGPDPNPVSIWLIVFGVWGVIVGIVILITGPIHDKKKKAKSGENP 780
OY 781 YASIDISKGENNGEONTDDVQTSF 805
OY 781 YASIDISKGENNGEONTDDVQTSF 805
Db 781 YASIDISKGENNGEONTDDVQTSF 805
Db 781 YASIDISKGENNGEONTDDVQTSF 805

```

```

RESULT 3
AAV72667 standard; protein: 805 AA.
AAV72667:
31-MAY-2001 (first entry)
Human angiotensin converting enzyme-2 (ACE-2).
Human: angiotensin converting enzyme-2; ACE-2; peptidyl dipeptidase A;
screening; therapy; hypertension; congestive heart failure; CHF;
inflammation; pain.
XX Homo sapiens.
OS
XX key Location/Qualifiers
XX Peptide 1..18
XX FT /label= Signal_peptide
XX FT 19..805
XX FT /label= Mature_ACE-2_protein

```

```

FT Domain 374..378
FT /label= ZBD
FT /note= "Zinc binding domain"
FT Domain 741..765
FT /label= TMD
FT /note= "Transmembrane domain; Hydrophobic region"
FT Domain 766..805
FT /label= Cytoplasmic_domain
US6194556-B1.
27-FEB-2001.
11-DEC-1997; 97US-0989299.
11-DEC-1997; 97US-0989299.
(MILL-) MILLENNIUM PHARM INC.
Acton SL, Robison KE;
WPI: 2001-210604/21.
N-PSDB: RAD02758.
Novel genes encoding angiotensin converting enzyme-2 useful as
antigenic or antigenic agents for therapeutics, diagnostics and
screening assays -
Claim 33; Fig 1; 76pp; English.
The present amino acid sequence is human angiotensin converting enzyme-2
(ACE-2), also referred as peptidyl dipeptidase A (EC 3.4.15.1). Nucleic
acid sequence encoding ACE-2 is useful as antisense or antigenic agents
for sequence specific modulation of gene expression or in the analysis of
single base-pair mutations in the gene. Nucleic acid sequence encoding
ACE-2 is useful in therapeutics, diagnostics and in screening assays.
ACE-2 antagonist is used to treat hypertension or congestive heart
failure (CHF). ACE agonist is used to reduce the inflammation and pain
resulting from an insect sting or bite, which was accompanied by an
injection of bradykinin. Anti-ACE-2 antibodies are used to monitor ACE-2
protein levels for determining the disease or condition associated with
an aberrant protein level.
Sequence 805 AA:
SQ
Query Match 100.0%; Score 4291; DB 22; Length 805;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFHNEADELFYQSSLASWNYNINITEENVQ 60
OY 1 MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFHNEADELFYQSSLASWNYNINITEENVQ 60
Db 1 MSSSSWLLSLVAVTAOSTIEBOAKTFIDKFHNEADELFYQSSLASWNYNINITEENVQ 60
OY 61 NMNNAQKMSAFLEKEOSTLAOMYPLQETIQNTLVKLOLQALQONGSSVLSDEKSKRLNTTL 120
OY 61 NMNNAQKMSAFLEKEOSTLAOMYPLQETIQNTLVKLOLQALQONGSSVLSDEKSKRLNTTL 120
Db 61 NMNNAQKMSAFLEKEOSTLAOMYPLQETIQNTLVKLOLQALQONGSSVLSDEKSKRLNTTL 120
OY 121 NTMSTIYSTGKVCNPDNPQECILLEPGLNEIMANSIDYERLMAWESMRSEVGKOLRPLY 180
OY 121 NTMSTIYSTGKVCNPDNPQECILLEPGLNEIMANSIDYERLMAWESMRSEVGKOLRPLY 180
Db 121 NTMSTIYSTGKVCNPDNPQECILLEPGLNEIMANSIDYERLMAWESMRSEVGKOLRPLY 180
OY 181 EEEVYLKNEKARAHNYEDYDGYWRGDEYVNGVDYDYSRGOLIEDVEHTFEELKPLYEHL 240
OY 181 EEEVYLKNEKARAHNYEDYDGYWRGDEYVNGVDYDYSRGOLIEDVEHTFEELKPLYEHL 240
Db 181 EEEVYLKNEKARAHNYEDYDGYWRGDEYVNGVDYDYSRGOLIEDVEHTFEELKPLYEHL 240
OY 241 HAYVRAKLMAVAPSYISPIGCLPAHLIDGMWGRFTNLYSLTVPFGOKPNIDVTDAVQ 300
OY 241 HAYVRAKLMAVAPSYISPIGCLPAHLIDGMWGRFTNLYSLTVPFGOKPNIDVTDAVQ 300
Db 241 HAYVRAKLMAVAPSYISPIGCLPAHLIDGMWGRFTNLYSLTVPFGOKPNIDVTDAVQ 300
OY 301 AMDAORIFKEAEKFEFVSGLPMMTQGFWENSMLTDPGNVOKAVCHPTAMDGKGFRLIM 360
OY 301 AMDAORIFKEAEKFEFVSGLPMMTQGFWENSMLTDPGNVOKAVCHPTAMDGKGFRLIM 360
Db 301 AMDAORIFKEAEKFEFVSGLPMMTQGFWENSMLTDPGNVOKAVCHPTAMDGKGFRLIM 360

```

QY 361 CTCTVMDPFLTAHHEMGIQYDMAYAAQFLLRNGANGFHEAVGEIMSLSAATPKHLKS 420  
 Db 361 CTCTVMDPFLTAHHEMGIQYDMAYAAQFLLRNGANGFHEAVGEIMSLSAATPKHLKS 420  
 QY 421 IGLSPDFQEDNETEINFLKQALITVGTLPPTMLEKRWKRVKGEIPKQOMKKWEM 480  
 Db 421 IGLSPDFQEDNETEINFLKQALITVGTLPPTMLEKRWKRVKGEIPKQOMKKWEM 480  
 QY 481 KREIVGVPEVPHDETYCDPASLFFHSNDYSFIRYRTYTLVQFOFQALCOAKHEBPLH 540  
 Db 481 KREIVGVPEVPHDETYCDPASLFFHSNDYSFIRYRTYTLVQFOFQALCOAKHEBPLH 540  
 QY 541 KCDISNSTEAGOKLFNMLRLGKSEPWTLALENVGAKMMVRPLNFEPLFTWLKQNK 600  
 Db 541 KCDISNSTEAGOKLFNMLRLGKSEPWTLALENVGAKMMVRPLNFEPLFTWLKQNK 600  
 QY 601 NSFVGMSTDMSPYADQSIKVRISLKSALGKAYEMNDENYLFSSVAVAMROYFLKVK 660  
 Db 601 NSFVGMSTDMSPYADQSIKVRISLKSALGKAYEMNDENYLFSSVAVAMROYFLKVK 660  
 QY 661 QMILFGEEDVAVANLKRISFNFFVAPKNVSDIIPREVEKAIKMSRSRINDAPFLNDN 720  
 Db 661 QMILFGEEDVAVANLKRISFNFFVAPKNVSDIIPREVEKAIKMSRSRINDAPFLNDN 720  
 QY 721 SLEFLGIQPTLGPNOFPVSIWLVGVYGVVIVLITGIRDKKKKKASGENP 780  
 Db 721 SLEFLGIQPTLGPNOFPVSIWLVGVYGVVIVLITGIRDKKKKKASGENP 780  
 QY 781 YASIDISKGNNGFQNTDVCYTSF 805  
 Db 781 YASIDISKGNNGFQNTDVCYTSF 805

RESULT 4  
 AAB48095  
 ID AAB48095 standard; Protein: 805 AA.  
 AC AAB48095;  
 DT 19-MAR-2001 (first entry)

DE Human Zace2 protein.  
 KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;  
 KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;  
 KW ventricular systolic dysfunction; renal impairment; heart failure;  
 KW scleroderma renal crisis; atherosclerosis; antinflammatory; human;  
 KW antitachytic; bradykinin inactivator.

OS Homo sapiens.  
 XX MO200070032-A1.  
 XX 23-NOV-2000.  
 XX 03-MAY-2000; 2000MO-US11932.  
 XX 13-MAY-1999; 99US-0311482.  
 XX 27-AUG-1999; 99US-0384706.  
 XX (ZIMO) ZYMOGENETICS INC.

PI Pladdington CS, Petrie CR, Shoemaker KE, Bishop PD;  
 DR WPI; 2001-025018/03.  
 DR N-PSDB; AAC84366, AAC84367.

PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory  
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases  
 XX associated with inflammation such as arthritis and enterocolitis -  
 PS Example 1; Page 95-100; 125pp; English.

CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-  
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood  
 CC pressure regulation and fertility. Zace2 can be expressed by standard  
 CC recombinant methodology. Zace2 polypeptides are useful for treating an  
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),  
 CC as targets for identifying new modulators of zinc protease activity, for  
 CC screening or identifying new modulators of zinc protease activity, for  
 CC inhibitors, and as a basis for rational drug design for inhibitory  
 CC molecules. The nucleic acids can be used to detect the expression of a  
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and  
 CC for detecting and localizing Zace2 gene expression in tissue samples,  
 CC to determine whether a subject's chromosomes contain a mutation in the  
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.  
 CC conditions, including left ventricular systolic dysfunction, progressive  
 CC renal impairment, scleroderma renal crisis, congestive heart failure due  
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be  
 CC used to treat infertility while Zace2 antagonists are used for inducing  
 CC infertility. The present sequence represents the human Zace2 protein.  
 XX

Query Match 805 AA:  
 Best Local Similarity 100.0%; Score 4291; DB 22; Length 805;  
 Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSIVAVTAOSTIEQAKTFDKRNHAEPLFOSSLASNNVTNTEENVQ 60  
 Db 1 MSSSSWLLSIVAVTAOSTIEQAKTFDKRNHAEPLFOSSLASNNVTNTEENVQ 60  
 QY 61 MNMAGDKWASAFLEQOSTLAQMYPLQETQNLVQLQALQOQSSVLSSEKSKRLTIL 120  
 Db 61 MNMAGDKWASAFLEQOSTLAQMYPLQETQNLVQLQALQOQSSVLSSEKSKRLTIL 120  
 QY 121 NTMSTIYSTGKVCNPDNPOECLLEPGNETMANSLDYNERLWAMESWSEVQKQRLPY 180  
 Db 121 NTMSTIYSTGKVCNPDNPOECLLEPGNETMANSLDYNERLWAMESWSEVQKQRLPY 180  
 QY 181 EBYVVLKNDMARANHYEDYGDYWRGDYEVNGVDYDSRGQLLEDVHPEELKPYEHL 240  
 Db 181 EBYVVLKNDMARANHYEDYGDYWRGDYEVNGVDYDSRGQLLEDVHPEELKPYEHL 240  
 QY 241 HAYYRAKIMNAPYISIPICGLPAHLGDMWGRWNLXSLYVFEQKKNIDVTDAVDQ 300  
 Db 241 HAYYRAKIMNAPYISIPICGLPAHLGDMWGRWNLXSLYVFEQKKNIDVTDAVDQ 300  
 QY 301 AMDAQRIFKEAEKFFVSVGLPNTQGFENSMULTDEGNVQKAVCHPTAMDLCGDFRILM 360  
 Db 301 AMDAQRIFKEAEKFFVSVGLPNTQGFENSMULTDEGNVQKAVCHPTAMDLCGDFRILM 360  
 QY 361 CTCTVMDPFLTAHHEMGIQYDMAYAAQFLLRNGANGFHEAVGEIMSLSAATPKHLKS 420  
 Db 361 CTCTVMDPFLTAHHEMGIQYDMAYAAQFLLRNGANGFHEAVGEIMSLSAATPKHLKS 420  
 QY 421 IGLSPDFQEDNETEINFLKQALITVGTLPPTMLEKRWKRVKGEIPKQOMKKWEM 480  
 Db 421 IGLSPDFQEDNETEINFLKQALITVGTLPPTMLEKRWKRVKGEIPKQOMKKWEM 480  
 QY 481 KREIVGVPEVPHDETYCDPASLFFHSNDYSFIRYRTYTLVQFOFQALCOAKHEBPLH 540  
 Db 481 KREIVGVPEVPHDETYCDPASLFFHSNDYSFIRYRTYTLVQFOFQALCOAKHEBPLH 540  
 QY 541 KCDISNSTEAGOKLFNMLRLGKSEPWTLALENVGAKMMVRPLNFEPLFTWLKQNK 600  
 Db 541 KCDISNSTEAGOKLFNMLRLGKSEPWTLALENVGAKMMVRPLNFEPLFTWLKQNK 600  
 QY 601 NSFVGMSTDMSPYADQSIKVRISLKSALGKAYEMNDENYLFSSVAVAMROYFLKVK 660  
 Db 601 NSFVGMSTDMSPYADQSIKVRISLKSALGKAYEMNDENYLFSSVAVAMROYFLKVK 660  
 QY 661 QMILFGEEDVAVANLKRISFNFFVAPKNVSDIIPREVEKAIKMSRSRINDAPFLNDN 720  
 Db 661 QMILFGEEDVAVANLKRISFNFFVAPKNVSDIIPREVEKAIKMSRSRINDAPFLNDN 720

Db 661 QMILFGEEDVAVANLKRISFNFTAPKAVSDIIPTEVEKAIKRSRKRINDAFRLNDN 720

Oy 721 SLEFLGIQPLTGPNNPPVSIWLVGVVGVIVGIVILITGIDRRKKKRRARSGENF 780

Db 721 SLEFLGIQPLTGPNNPPVSIWLVGVVGVIVGIVILITGIDRRKKKRRARSGENF 780

Oy 781 YASIDISKGENNPGFNTDVOVSF 805

Db 781 YASIDISKGENNPGFNTDVOVSF 805

RESULT 5

AAU99701 ID AAU99701 standard; protein; 805 AA.

XX AAU99701;

XX 24-SEP-2002 (first entry)

DE Human angiotensin converting enzyme-2 (ACE-2) protein.

XX Human; angiotensin converting enzyme-2; ACE-2; body weight disorder;

KW muscle mass; body fat; obesity; diabetes; atherosclerosis; weight loss;

KW lipid metabolism; weight gain; anorexia; cachexia; bulimia; sepsis;

KW familial partial lipodystrophy; hypercholesterolemia; hyperlipidaemia;

KW aberrant metabolic rate; heart failure; left ventricular hypertrophy;

KW neurodegenerative disorder; peptide hormone; cytokine processing;

KW myocardial infarction; cardiomyopathy; inflammatory bowel disease;

KW systemic inflammation response syndrome; polytrauma; pain; stroke;

KW bone destruction; rheumatoid arthritis; osteoarthritis; asthma;

KW periodontal disease; dysmenorrhoea; premature labour; brain oedema;

KW focal injury; diffuse axonal injury; reperfusion injury; scar formation;

KW cerebral vasospasm; subarachnoid haemorrhage; allergic disorder;

KW adult respiratory distress syndrome; wound healing; appetite;

KW body mass index.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Peptide 1..18

FT /label= Signal\_peptide

FT 19..805

FT /label= Mature\_human\_ACE\_2\_protein

FT Protein

XX WO200239997-A2.

XX 23-MAY-2002.

XX 31-OCT-2001; 2001WO-US45703.

XX 01-NOV-2000; 2000US-0704216

XX 29-MAY-2001; 2001US-0870382

XX 19-OCT-2001; 2001US-371741P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Action SL, Ocala TD, Gould AE, Dales NA, Guan B, Brown JA;

PI Palane M, Kadambi VJ, Solomon M, Stricker-Krongrad A;

XX N-PSDB; ABR67623.

DR WPI: 2002-547572/58.

XX Treating body weight disorder and increasing muscle mass comprises

PT administering angiotensin converting enzyme-2 modulating compound

XX Example 5; Page 387-390; 395pp; English.

XX The present invention describes a new method of treating a body weight

CC disorder, increasing muscle mass and decreasing body fat by

CC administration of angiotensin converting enzyme (ACE)-2 modulating

CC compound. The invention can be used for treating body weight disorders,

CC particularly obesity of at least grade 1, diabetes, atherosclerosis and

CC a state associated with lipid metabolism. The method is used for treating

CC rapid weight loss, rapid weight gain, anorexia, cachexia, bulimia,

CC generalised partial lipodystrophy, familial partial lipodystrophy,

CC hypercholesterolemia, hyperlipidaemia, an aberrant metabolic rate,

CC congestive heart failure, chronic heart failure, left ventricular

CC hypertrophy, acute heart failure, neurodegenerative disorders (e.g.

CC Alzheimer's disease, Parkinson's disease and Huntington's disease),

CC diseases associated with peptide hormones or cytokine processing,

CC myocardial infarction, cardiomyopathy, systemic inflammation response

CC syndrome, sepsis, polytrauma, inflammatory bowel disease, acute and

CC chronic pain, bone destruction in rheumatoid arthritis and osteoarthritis

CC and periodontal disease, dysmenorrhoea, premature labour, brain oedema

CC following focal injury, diffuse axonal injury, stroke, reperfusion

CC injury, cerebral vasospasm after subarachnoid haemorrhage, allergic

CC disorders including asthma, adult respiratory distress syndrome, wound

CC healing and scar formation. The invention decreases the appetite,

CC increases muscle mass and decreases body fat of subject having body mass

CC index of greater than 23 (preferably 24.9)kg/m<sup>2</sup>. The present amino

CC acid sequence represents the human ACE-2 protein of the invention.

XX

Sequence 805 AA:

Query Match 100.0%; Score 4291; DB 23; Length 805;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 805; Conservative 0;

Oy 1 MSSSSMLLSLVAVTAOSTIEBOAKTFLDKFHEADELFYOSLSAMNYNTNTEENYQ 60

Db 1 MSSSSMLLSLVAVTAOSTIEBOAKTFLDKFHEADELFYOSLSAMNYNTNTEENYQ 60

Oy 61 MNMAGDKMSAFLEKOSTLAQMYPLQEIQNTLVKIQLOALQONGSSVSEDKSKRLNTLL 120

Db 61 MNMAGDKMSAFLEKOSTLAQMYPLQEIQNTLVKIQLOALQONGSSVSEDKSKRLNTLL 120

Oy 121 NTSTSTSTKVCNPNPPOCLLEPGCINPMANSIDYNERLAWMSSEVKOLRPLY 180

Db 121 NTSTSTSTKVCNPNPPOCLLEPGCINPMANSIDYNERLAWMSSEVKOLRPLY 180

Oy 121 NTSTSTSTKVCNPNPPOCLLEPGCINPMANSIDYNERLAWMSSEVKOLRPLY 180

Db 121 NTSTSTSTKVCNPNPPOCLLEPGCINPMANSIDYNERLAWMSSEVKOLRPLY 180

Oy 181 EEEYVLKNEARANHEDYDYGWRCDEYGVGYDSRGOLLEDEYHTEETKPLYEH 240

Db 181 EEEYVLKNEARANHEDYDYGWRCDEYGVGYDSRGOLLEDEYHTEETKPLYEH 240

Oy 181 EEEYVLKNEARANHEDYDYGWRCDEYGVGYDSRGOLLEDEYHTEETKPLYEH 240

Db 181 EEEYVLKNEARANHEDYDYGWRCDEYGVGYDSRGOLLEDEYHTEETKPLYEH 240

Oy 241 HAYVRAKIMNAYPSYISPIGLPAHLIGDMGFRWNLVSLYVPSGCKRPIVDTAMVQ 300

Db 241 HAYVRAKIMNAYPSYISPIGLPAHLIGDMGFRWNLVSLYVPSGCKRPIVDTAMVQ 300

Oy 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWENSMLTDPENYQKAVCHPTAMDIGKDRILM 360

Db 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWENSMLTDPENYQKAVCHPTAMDIGKDRILM 360

Oy 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWENSMLTDPENYQKAVCHPTAMDIGKDRILM 360

Db 301 AMDAQRIFKEAEKFFVSGLPNMTQGFWENSMLTDPENYQKAVCHPTAMDIGKDRILM 360

Oy 361 CTKYTMDDFLTAHHEMGIQDYMAVAAQPLRLNGANEGFHEAVGEIMLSAATPKHLKS 420

Db 361 CTKYTMDDFLTAHHEMGIQDYMAVAAQPLRLNGANEGFHEAVGEIMLSAATPKHLKS 420

Oy 361 CTKYTMDDFLTAHHEMGIQDYMAVAAQPLRLNGANEGFHEAVGEIMLSAATPKHLKS 420

Db 361 CTKYTMDDFLTAHHEMGIQDYMAVAAQPLRLNGANEGFHEAVGEIMLSAATPKHLKS 420

Oy 421 IGLSPDFQEDNTEINFLKQALITVGLPFTYMLEKRWNYFKGIPRQDMKKWEM 480

Db 421 IGLSPDFQEDNTEINFLKQALITVGLPFTYMLEKRWNYFKGIPRQDMKKWEM 480

Oy 421 IGLSPDFQEDNTEINFLKQALITVGLPFTYMLEKRWNYFKGIPRQDMKKWEM 480

Db 421 IGLSPDFQEDNTEINFLKQALITVGLPFTYMLEKRWNYFKGIPRQDMKKWEM 480

Oy 481 KREIVGVVEPYPHDETCDPASLPHVSNDSFIRYRTLYXOPFOALQAAKHESPLH 540

Db 481 KREIVGVVEPYPHDETCDPASLPHVSNDSFIRYRTLYXOPFOALQAAKHESPLH 540

Oy 481 KREIVGVVEPYPHDETCDPASLPHVSNDSFIRYRTLYXOPFOALQAAKHESPLH 540

Db 481 KREIVGVVEPYPHDETCDPASLPHVSNDSFIRYRTLYXOPFOALQAAKHESPLH 540

Oy 541 KCDISNTEAGOKLFNMLRLKSEBPTLALENVGAKNMVRLPLNFEPLFTWLKQDNK 600

Db 541 KCDISNTEAGOKLFNMLRLKSEBPTLALENVGAKNMVRLPLNFEPLFTWLKQDNK 600

Oy 541 KCDISNTEAGOKLFNMLRLKSEBPTLALENVGAKNMVRLPLNFEPLFTWLKQDNK 600

Db 541 KCDISNTEAGOKLFNMLRLKSEBPTLALENVGAKNMVRLPLNFEPLFTWLKQDNK 600

Oy 601 NSFWGWSWDMSPYADQSIKVRISLKSALGDKAYEMNDENMLTFRSSVAYAMROFLVKVN 660

Db 601 NSFWGWSWDMSPYADQSIKVRISLKSALGDKAYEMNDENMLTFRSSVAYAMROFLVKVN 660

Oy 601 NSFWGWSWDMSPYADQSIKVRISLKSALGDKAYEMNDENMLTFRSSVAYAMROFLVKVN 660

Db 601 NSFWGWSWDMSPYADQSIKVRISLKSALGDKAYEMNDENMLTFRSSVAYAMROFLVKVN 660

Oy 661 QMILFGEEDVAVANLKRISFNFTAPKAVSDIIPTEVEKAIKRSRKRINDAFRLNDN 720

Db 661 QMILFGEEDVAVANLKRISFNFTAPKAVSDIIPTEVEKAIKRSRKRINDAFRLNDN 720



QY 721 SLEFLGIQPTLGPNNPVSIMLIVGVGVVGLVILITFTGIRDRKKKKKARSGENP 780  
 DB 721 SLEFLGIQPTLGPNNPVSIMLIVGVGVVGLVILITFTGIRDRKKKKKARSGENP 780  
 QY 781 YASIDISKGNNPGFQNTDVOYTSF 805  
 DB 781 YASIDISKGNNPGFQNTDVOYTSF 805

RESULT 6  
 AAE20353  
 ID AAE20353 standard; Protein; 805 AA;  
 AC AAE20353;  
 XX  
 XX 18-JUN-2002 (first entry)  
 DE Human ACE-2 full-length protein.  
 XX  
 KW Human; angiotensin converting enzyme-2; ACE-2; therapy; hypertension;  
 KW peptidyl dipeptidase A; blood pressure; hypotension; atherosclerosis;  
 KW myocardial infarction; heart failure; arrhythmia; renal failure; gene;  
 KW inflammation; fertility; enzyme; EC 3.4.15.1.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Peptide  
 FT /label= Signal\_peptide  
 FT Protein  
 FT /note= "Mature ACE-2 protein"  
 FT Domain  
 FT /note= "Extracellular domain"  
 FT /note= "Zinc binding domain (ZBD)"  
 FT Domain  
 FT /note= "Transmembrane domain"  
 FT /note= "766..805"  
 FT /note= "Cytoplasmic domain"  
 FT  
 XX WO200212471-A2.  
 PN  
 XX 14-FEB-2002.  
 PD  
 XX 09-AUG-2001; 2001WO-US25059.  
 PE  
 XX 09-AUG-2000; 2000US-0635501.  
 PR  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 XX  
 PI Acton S, Robison KE, Hsieh FY;  
 XX  
 DR WPI; 2002-257481/30.  
 DR N-PSDB; AAD32586.  
 DR  
 XX  
 PT Isolated human polypeptide, known as angiotensin converting enzyme-2,  
 PT useful for treating or preventing the development of an abnormal blood  
 PT pressure or related diseases, e.g. hypertension, heart failure or  
 PT myocardial infarction -  
 PT  
 XX  
 XX Claim 2; Fig 1; 218P; English.  
 PS  
 XX  
 CC The invention relates to human angiotensin converting enzyme-2 (ACE-2)  
 CC polypeptides and polynucleotides. ACE-2 is also known as peptidyl  
 CC dipeptidase A (EC 3.4.15.1). Polypeptides of the invention are useful  
 CC for treating or preventing the development of abnormal blood pressure  
 CC and diseases or disorders associated with the protein in a subject. The  
 CC diseases include hypertension, hypotension, congestive heart failure,  
 CC chronic heart failure, acute heart failure, myocardial infarction,  
 CC atherosclerosis, arrhythmia and renal failure. They are also useful  
 CC for treating inflammatory conditions and diseases relating to fertility.  
 CC The present sequence is human full-length ACE-2 protein.  
 CC  
 XX

SQ Sequence 805 AA;  
 Query Match 100.0%; Score 4291; DB 23; Length 805;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSLVAVTAOSTIEBOAKTFLDKFNHEADLFYQSSLASMNTNTTEBNVQ 60  
 DB 1 MSSSSWLLSLVAVTAOSTIEBOAKTFLDKFNHEADLFYQSSLASMNTNTTEBNVQ 60  
 QY 61 NNNNAGDKWSAFLEKQSTLAQMTPLQETLQNLTVLQLQALQONSSVLSDEKSKRLTIL 120  
 DB 61 NNNNAGDKWSAFLEKQSTLAQMTPLQETLQNLTVLQLQALQONSSVLSDEKSKRLTIL 120  
 QY 121 NTMSTIYSTGKVCNPDNPOECLELLEPGINEIMANSIDYNERLWAMESRSEVQKQLRPLY 180  
 DB 121 NTMSTIYSTGKVCNPDNPOECLELLEPGINEIMANSIDYNERLWAMESRSEVQKQLRPLY 180  
 QY 181 EEEVVLKEMARAHNYEDYGYRWGDIYEVNGVDGYDSRKQLLEDVETHEEIKPLYEHL 240  
 DB 181 EEEVVLKEMARAHNYEDYGYRWGDIYEVNGVDGYDSRKQLLEDVETHEEIKPLYEHL 240  
 QY 241 HAYVRAKIMNAPPSYISPIGCLPAHLGDMWGRFTNLYSILYVPEGOKPVIDVTDAWVQ 300  
 DB 241 HAYVRAKIMNAPPSYISPIGCLPAHLGDMWGRFTNLYSILYVPEGOKPVIDVTDAWVQ 300  
 QY 301 AMDAQRIFKAEKEFFVSGVGLPNTQGFWEWSMLTDPGNVOKAVCHPTAMDIGKDFRILM 360  
 DB 301 AMDAQRIFKAEKEFFVSGVGLPNTQGFWEWSMLTDPGNVOKAVCHPTAMDIGKDFRILM 360  
 QY 361 CTKVTMDDFLTAHHEMHIQYDMAVYAAQPLFLNKGNEGFHEAVGFTMSLSTATPRHLKS 420  
 DB 361 CTKVTMDDFLTAHHEMHIQYDMAVYAAQPLFLNKGNEGFHEAVGFTMSLSTATPRHLKS 420  
 QY 421 IGLISPDFOEDNETEINFLKQALITVGLPFTYMLEKRWVYFKEIIPKDDMKKRWEM 480  
 DB 421 IGLISPDFOEDNETEINFLKQALITVGLPFTYMLEKRWVYFKEIIPKDDMKKRWEM 480  
 QY 481 KREIVGVPEVPDHEVCDPASLPHVSNDSYSTRYTRTLXQFOEALCOAKHEGPIH 540  
 DB 481 KREIVGVPEVPDHEVCDPASLPHVSNDSYSTRYTRTLXQFOEALCOAKHEGPIH 540  
 QY 541 KCDISNSTEAGOKLFNNLRGKSEPTLALENVGAKNNVRLPLNYPPLTFLTKDOK 600  
 DB 541 KCDISNSTEAGOKLFNNLRGKSEPTLALENVGAKNNVRLPLNYPPLTFLTKDOK 600  
 QY 601 NSFVGMSTWSPYADOSIKVRISLKSALGDKAYEMNDNEMYLFRSSVAYAMROYFLKVN 660  
 DB 601 NSFVGMSTWSPYADOSIKVRISLKSALGDKAYEMNDNEMYLFRSSVAYAMROYFLKVN 660  
 QY 661 QMILFGEEDVRVANKPRISFNFVAPKNSDILIPREVEKAIKRSRINDARLNDN 720  
 DB 661 QMILFGEEDVRVANKPRISFNFVAPKNSDILIPREVEKAIKRSRINDARLNDN 720  
 QY 721 SLEFLGIQPTLGPNNPVSIMLIVGVGVVGLVILITFTGIRDRKKKKKARSGENP 780  
 DB 721 SLEFLGIQPTLGPNNPVSIMLIVGVGVVGLVILITFTGIRDRKKKKKARSGENP 780  
 QY 781 YASIDISKGNNPGFQNTDVOYTSF 805  
 DB 781 YASIDISKGNNPGFQNTDVOYTSF 805

RESULT 7  
 AAU09092  
 ID AAU09092 standard; Protein; 711 AA;  
 AC AAU09092;  
 XX  
 XX 20-DEC-2001 (first entry)  
 DE Novel human protein NHP #1.  
 XX

Human: novel human protein: NHP; antidiabetic; antirheumatic;  
 neuroprotective; cytoskeletal; antilipidemic; antileukemic;  
 antihuman immunodeficiency virus; antistaphylococcal; vasotropic; cardiant;  
 hypotensive; anorectic; antileukemic; neuroleptic; anticonvulsant;  
 antitumor; immunosuppressive; cerebroprotective; antimicrobial;  
 antidiabetic; antileukemic; antiparasitic; thyromimetic;  
 immunomodulatory; antileukemic; dermatological; vasoconstrictor;  
 gastrointestinal disorder; cardiovascular disorder; hypertension;  
 coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;  
 coxidia; male infertility; impotence; testicular cancer; lung tumor;  
 hyperproliferative disorder; pulmonary system disorder;  
 central nervous system disorder; bone disorder; Parkinson's disease;  
 neurodegenerative disease; Alzheimer's disease; dementia; paraneoplasia;  
 Huntington's disease; schizophrenia; amyotrophic lateral sclerosis;  
 panic disorder; learning disability; immune system disorder;  
 psychosis; autism; sleep disorder; immune system disorder;  
 Hashimoto's thyroiditis; musculo-skeletal system disorders;  
 multiple sclerosis; ischemic brain injury; stroke; infectious disease;  
 diabetes mellitus; immunological disorder; asthma; AIDS; immunogen;  
 acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;  
 inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;  
 neural system disorder; respiratory disorder; olfactory disorder;  
 wound healing; chromosome X.

Homo sapiens.

Key	Location/Qualifiers
Domain	1..681 Extracellular_domain
Region	48..55 Immunogenic-epitope
Region	110..118 Immunogenic-epitope
Region	136..146 Immunogenic-epitope
Region	151..158 Immunogenic-epitope
Misc-difference	219
Label= OTHER	
/note= "Other- Any amino acid encoded by WST"	
Misc-difference	240
Label= OTHER	
/note= "Other- Any amino acid encoded by RCC"	
Misc-difference	499
Label= OTHER	
/note= "Other- Any amino acid encoded by NTT"	
Domain	682..698
Label= Transmembrane_domain	

WO200174896-A1.

11-OCT-2001.

02-APR-2001: 2001WO-US10542.

03-APR-2000: 2000US-194118P.

29-SEP-2000: 2000US-236384P.

(HUMA-) HUMAN GENOME SCI INC.

Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;  
 Li Y, Dillon PJ;  
 WPI: 2001-626394/72.  
 N-PSDB; AAS14880.

New human proteins, useful for diagnosing, treating, preventing and/or  
 prognosing disorders related to the proteins, including cardiovascular  
 disorders, autoimmune disorders and reproductive disorders

Claim 11, Page 298-301; 318pp; English.

The invention relates to novel human proteins (NHP) and the  
 nucleic acids that encode them and antibodies raised against them.  
 The proteins, antibodies and nucleic acids are useful in the diagnosis,  
 prognosis, prevention and/or treatment of diseases and/or disorders  
 involving vasoconstriction, gastrointestinal disorders, cardiovascular  
 disorders (e.g. hypertension, erectile dysfunction, high blood pressure,  
 coronary heart disease and arteriosclerosis), anorexia, obesity, bulimia,  
 coxidia, disorders of small intestine, disorders of reproductive system,  
 (e.g. male infertility and/or impotence), disorders of pulmonary system,  
 and other hyperproliferative disorders, bone disorders, neurodegenerative  
 central nervous system disorders, bone disorders, dementia, paraneoplasia,  
 disease, Huntington's disease, schizophrenia, amyotrophic lateral sclerosis,  
 panic disorder, learning disabilities, immune system disorders (e.g.  
 psychoses, autism, sleep disorders), renal and musculo-skeletal system disorders,  
 Hashimoto's thyroiditis), renal and musculo-skeletal system disorders,  
 central nervous system disorders (e.g. multiple sclerosis, ischemic  
 brain injury and/or stroke), infectious diseases, diabetes mellitus,  
 immunological disorders (e.g. asthma, acquired immunodeficient syndrome  
 (AIDS), leukaemia, rheumatoid arthritis, inflammatory bowel disease,  
 sepsis, acne, psoriasis and lupus erythematosus), neural system  
 disorders, respiratory disorders, olfactory disorders and wound  
 healing. The present sequence represents an NHP of the invention the  
 gene for which is located on the X chromosome.

Sequence 711 AA;

Query Match 88.04; Score 3775; DB 22; Length 711;  
 Best Local Similarity 99.35; Pred. No. 1.3e-311; 32X  
 Matches 704; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY	62	MNNAGDKWSAFLEKREOSTLAQWYPIQELQNLVVKLOLQALQONSSVLSSEKSKRLNTIIN	121
DB	1	MNNAGDKWSAFLEKREOSTLAQWYPIQELQNLVVKLOLQALQONSSVLSSEKSKRLNTIIN	60
QY	122	TWSTIYSGKYCNPDPNPECELLLEBGLNETMANSIDYNERIMAMESRSPGOLRPIYE	181
DB	61	TWSTIYSGKYCNPDPNPECELLLEBGLNETMANSIDYNERIMAMESRSPGOLRPIYE	120
QY	182	ERVVLKNEKARAHNEYEDYGDYWRGDEYVNGVDYDYSRGOLLEDEVFEEIRPLEYELH	241
DB	121	ERVVLKNEKARAHNEYEDYGDYWRGDEYVNGVDYDYSRGOLLEDEVFEEIRPLEYELH	180
QY	242	AVYRAKLNAYSPYISPIGCLPAHLGDMGREFWNTLYSLTPPGQKPNIDYDAMVDA	301
DB	181	AVYRAKLNAYSPYISPIGCLPAHLGDMGREFWNTLYSLTPPGQKPNIDYDAMVDA	240
QY	302	WDAQRTEKFAEKFPVSVGLPNTGQFENSMITDGNVQKAVCHPTAMDLSKGFRTIMC	361
DB	241	WDAQRTEKFAEKFPVSVGLPNTGQFENSMITDGNVQKAVCHPTAMDLSKGFRTIMC	300
QY	362	TKVTNDPFLAHHEMHGHIQDMAVAAOPFLFNGANEGFHEAVGEIMSLSATPRKHLSI	421
DB	301	TKVTNDPFLAHHEMHGHIQDMAVAAOPFLFNGANEGFHEAVGEIMSLSATPRKHLSI	360
QY	422	GLSPDFQEDNETEINFLKQALTYVGTLPPTMYLEKRMWYFKETPKDQMKKWMK	481
DB	361	GLSPDFQEDNETEINFLKQALTYVGTLPPTMYLEKRMWYFKETPKDQMKKWMK	420
QY	482	RELTVGVEPVPDHEVCDPASLFHNSNDYSFRTYTRTLYQFOFOEALQQAHEGRLHK	541
DB	421	RELTVGVEPVPDHEVCDPASLFHNSNDYSFRTYTRTLYQFOFOEALQQAHEGRLHK	480
QY	542	CDISNSTEAGOKLFNMLRQKSEPTWLALENVYGAKNMVRPLNLFEPDLTKLQDKN	601
DB	481	CDISNSTEAGOKLFNMLRQKSEPTWLALENVYGAKNMVRPLNLFEPDLTKLQDKN	540
QY	602	SFVGSKTDSPYADQSTKVAISLSKALGKAVEMNDENYLRSSVAAVAMROFELKVNQ	661
DB	541	SFVGSKTDSPYADQSTKVAISLSKALGKAVEMNDENYLRSSVAAVAMROFELKVNQ	600
QY	662	MILFGEEDVAVANLKPRLISENFVTPAKNVSDIIPTEVEKAKIRSRSRINDAPLANS	721

CC Zace2 gene in a biological sample, as probes for *in vivo* diagnosis and  
CC for detecting and localizing Zace2 gene expression in tissue samples',  
CC to determine whether a subject's chromosomes contain a mutation in the  
CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.  
CC Inhibitors of ACE are used for treating hypertension of various  
CC conditions, including left ventricular systolic dysfunction, progressive  
CC renal impairment, scleroderma renal crisis, congestive heart failure due  
CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be  
CC used to treat infertility while Zace2 antagonists are used for inducing  
CC infertility. The present sequence represents the mouse Zace2-5 protein.  
XX  
XX Sequence 805 AA:  
XX

Query Match	83.4%	Score 3579	DB 22	Length 805
Best Local Similarity	82.18	Pred. No. 6,9e-295		
Matches	661	Conservative 60	Mismatches 84	Indels 0
Gaps	0			
QY	1	MSSSWLLSLVATAASTTEEDQAKTEFLDKFNHEADLLPYGSSLSMWNNTNTTEBNQ	60	
DB	1	MSSSWLLSLVAVTAQTQSLTEENAKTFLLNPNQEAEDLSYGSSLSMWNNTNTTEBNQ	60	
QY	61	MNNAAGKWSAFKEOSTLQMPLOETQNLTKVLQLOALQNGSSVLSEDSKRRLNTL	120	
DB	61	KMSLAAKWSAFKEEOSTKQASFSLOEIPIIRKLOALQLOOGSSALADKKNQNTL	120	
QY	121	NTMSTIYSTGKVCNPDQECILLEPLINIMANSIDYERLWAMSSEVSGKQRLPX	180	
DB	121	NTMSTIYSTGKVCNPKQECILLEPLIDELIMATSTDYNSRLWAMSGRAEYEGKQRLPX	180	
QY	181	EELYVLKNEMARANHYEDYGDYRGDYEYNGVDGYDYSKGQLLIEDEYHTFEETKPLYEHL	240	
DB	181	EELYVLKNEMARANNYNDYGDYRGDYEAGDGYNYNKNQOLLEDEYRTFAETKPLYEHL	240	
QY	241	HAYVRAKIMANAYPSYISPLICLPAHLILGDMGGEFNNILSLVPPGQRKNDIVTAMNQ	300	
DB	241	HAYVRKRLMDTYSYISPLICLPAHLILGDMGGEFNNILPLTVPAQKNDIVTAMNQ	300	
QY	301	AMNARLFRKAEKFEYVSVGLPNNTQGFWMNSMLTDFGNVQKAVCHPTAMDLSKGFRLIM	360	
DB	301	GMDAERLEDEAEKFEYVSVGLPHNTQGFWMNSMLTEPDAKRVYCHPTAMDLSKGFRLIM	360	
QY	361	CTVYTMDDFETLHHEGHIQYDMAVAAPQELLRNANGSEFHAAGEIMSLSAATPKHLKS	420	
DB	361	CTVYTMDDNPLTAHHEGHIQYDMAVAAPQELLRNANGSEFHAAGEIMSLSAATPKHLKS	420	
QY	421	IGLLSPDEQEDNETEINFLKQALLTVGLTPTYMLEKRMVVFGEKLPKQOMKKWEM	480	
DB	421	IGLLSPDEQEDSETEINFLKQALLTVGLTPTYMLEKRMVVFGEKLPKQOMKKWEM	480	
QY	481	KRIYVAVPVPHDEYCDPRLAFVNSDYSTRYTTTLVYQFOFQALCOAARHEGRL	540	
DB	481	KRIYVAVPRLPHDEYCDPRLAFVNSDYSTRYTTTLVYQFOFQALCOAARVNSLH	540	
QY	541	KCDISNSTEAGOKLRLNMLRKSGSEFWTJALEYVVGAKMANNRPLINYEPLFTLKOONK	600	
DB	541	KCDISNSTEAGOKLRLNMLSLGSEFWTJALEYVVGAKMANNRPLINYEPLFTLKOONK	600	
QY	601	NSFVQWSTDMSYTAOOSTIKVRLSKLSALGDKRAYEMDNENYLFSSVAAYAMROYETLVKN	660	
DB	601	NSFVQWSTDMSYTAOOSTIKVRLSKLSALGANAYEMTNENYLFSSVAAYAMRKYSILKN	660	
QY	661	QMLTEGEDVAVRANKLRPISFNFTYAPKANNSDILPTVEYKALRMSRSTINDAFRLND	720	
DB	661	QVPLLEEDVAVRANKLRPISFNFTYAPKANNSDVIRSEVEDALRMSRGRINDFVGLND	720	
QY	721	SLEFLGIQPLTGPNQPVSLIMLVFGVWGVIVVGGVILITFGIRPKRKKRARSQEND	780	
DB	721	SLEFLGIQPLTEPPQPVSLIMLVFGVWVAVVGVILITFGIRPKRKKRNETKREND	780	
QY	781	YASIDISGENNPPQNTDVOYTSF	805	
DB	781	YDSMDIGESNAGFONSDAQTSE	805	

XX  
XX

XX 24-OCT-2001 (first entry)  
 DT  
 XX  
 DE Human PRO1885 polypeptide sequence.  
 XX  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIa; gene therapy.  
 OS  
 XX Homo sapiens.  
 XX  
 PN WO200140466-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 01-DEC-2000; 2000WO-US32678.  
 XX  
 PR 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 09-DEC-1999; 99US-0170262.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04414.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 20-MAR-2000; 2000WO-US05601.  
 PR 20-MAR-2000; 2000WO-US07377.  
 PR 21-MAR-2000; 2000WO-US07532.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 10-NOV-2000; 2000WO-US30873.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX  
 DR WPI; 2001-408281/43.  
 DR N-PSDB; AAS21279.  
 XX  
 XX  
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical  
 XX  
 PS Claim 12; Fig 72; 813pp; English.  
 XX  
 XX AA012172-AA012446 represent novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from

CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC of T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIa. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 XX  
 SQ Sequence 555 AA;

Query Match 69.4%; Score 2979; DB 22; Length 555;  
 Best Local Similarity 99.8%; Pred. No. 3.8e-244;  
 Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSLVAVTAOSTTEEDAKTLDKFNHEADLFYQSSLSAMWNTNTEENQ 60  
 DB 1 MSSSSWLLSLVAVTAOSTTEEDAKTLDKFNHEADLFYQSSLSAMWNTNTEENQ 60  
 QY 61 MNNAAGKMSAFLEKOSTLAOMYPLQEIOMLYKLDLQDQNGSSVLEDSKRLNTIL 120  
 DB 61 MNNAAGKMSAFLEKOSTLAOMYPLQEIOMLYKLDLQDQNGSSVLEDSKRLNTIL 120  
 QY 121 NTMSTIYSTGKVCNPDNPOECLELLEPGLNEIMANSIDYNERLAMEMSRSEVGKOLRPLY 180  
 DB 121 NTMSTIYSTGKVCNPDNPOECLELLEPGLNEIMANSIDYNERLAMEMSRSEVGKOLRPLY 180  
 QY 181 EBYVVLKNEMARANHYEDGDYKRGDYEVNGVDGIDYSRGQLEDEHTEFEIKPLYEHL 240  
 DB 181 EBYVVLKNEMARANHYEDGDYKRGDYEVNGVDGIDYSRGQLEDEHTEFEIKPLYEHL 240  
 QY 241 HAVYRAKLMAAYPSYISPIGCLPAHLGDMGREFNTNLSLTPRGCKRNTDYDAWDQ 300  
 DB 241 HAVYRAKLMAAYPSYISPIGCLPAHLGDMGREFNTNLSLTPRGCKRNTDYDAWDQ 300  
 QY 301 AMDAQRIFEAKRFYVSVGLPMNTQGEWNSMLTDPGVQKAVCPHTAMDGKGFRTILM 360  
 DB 301 AMDAQRIFEAKRFYVSVGLPMNTQGEWNSMLTDPGVQKAVCPHTAMDGKGFRTILM 360  
 QY 361 CTKVTMDDELTAHHEMGHIOYDMAAOPFLIRNGANGFHEANGELMSLSAATPKHLKS 420  
 DB 361 CTKVTMDDELTAHHEMGHIOYDMAAOPFLIRNGANGFHEANGELMSLSAATPKHLKS 420  
 QY 421 IGLSPDFOEDNETEINFLKQALTYGTLPTMYLEKRWMYFKGELIPKDDMMKKWMEK 480  
 DB 421 IGLSPDFOEDNETEINFLKQALTYGTLPTMYLEKRWMYFKGELIPKDDMMKKWMEK 480  
 QY 481 KREIVGVEPVPHDETYCDPASLFHVSNDYSFITYRTLYXQFQFALCQAAKHEGPIH 540  
 DB 481 KREIVGVEPVPHDETYCDPASLFHVSNDYSFITYRTLYXQFQFALCQAAKHEGPIH 540  
 QY 541 KCDISNSTEAGQKL 554  
 DB 541 KCDISNSTEAGQKL 554

RESULT 11  
 AA67311  
 ID AA67311 standard; protein; 480 AA.  
 XX  
 AC AA67311;  
 XX  
 DT 11-Apr-2000 (first entry)  
 XX  
 DE Human MPRO15 amino acid sequence #2.  
 XX  
 KW MPRO15; treatment; hypertension; human; myocardial disease; apoplexy;  
 KW heart disease; apoplexy; heart disease; nervous denaturation; hormone;  
 KW Alzheimer's disease; cytokine.  
 OS  
 XX Homo sapiens.  
 XX

PN JP11318472-A.  
 KW 24-NOV-1999.  
 XX 22-JAN-1999; 99GP-0014949.  
 XX 13-MAY-1998; 98GB-0010373.  
 PR 18-AUG-1998; 98GB-0018009.  
 XX (SMIK ) SMITHKLINE BEECHAM PLC.  
 PA WPI: 2000-109268/10.  
 DR MPROT15 polypeptide and MPROT15 polynucleotides - useful for the  
 XX treatment of hypertension, myocardial diseases, apoplexy, heart  
 PT diseases, nervous denaturation, Alzheimer's disease etc.  
 PS Claim 19; Page 20-21; 22pp; Japanese.  
 CC This is amino acid sequence #2 of human MPROT15. The MPROT15  
 CC polynucleotide and polypeptide sequences can be used for the treatment of  
 CC hypertension, myocardial diseases, apoplexy, heart diseases, nervous  
 CC denaturation, Alzheimer's disease and diseases related to the processing  
 CC of peptide hormones and cytokines.  
 CC  
 SQ Sequence 480 AA:  
 Query Match 59.2%; Score 2539; DB 21; Length 480;  
 Best local Similarity 100.0%; Pred. No. 7.4e-207;  
 Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 LVAVTAAGSTIEHOAKTFEDKFNHAEADLFYQSSLSASWNTNTTEENVQNMNAGDKMS 70  
 Db 10 LVAVTAAGSTIEHOAKTFEDKFNHAEADLFYQSSLSASWNTNTTEENVQNMNAGDKMS 69  
 QY 71 AFLKEOSTLAQMTPLQEIOMLVKLOLQLOALQONGSSVISEKSKRLNTLTMNSTIYSTG 130  
 Db 70 AFLKEOSTLAQMTPLQEIOMLVKLOLQLOALQONGSSVISEKSKRLNTLTMNSTIYSTG 129  
 QY 131 KVCMPDNPQECLEPELNLGELNEMSLDYNERLWAMESRSEVKOLRPLYERYVYLKNEK 190  
 Db 130 KVCMPDNPQECLEPELNLGELNEMSLDYNERLWAMESRSEVKOLRPLYERYVYLKNEK 189  
 QY 191 ARANHEDEYDGYWNGDEYVNGVDGYDSRQQLIEDYEHFEEFKKPLYELHAYVRAKLMN 250  
 Db 190 ARANHEDEYDGYWNGDEYVNGVDGYDSRQQLIEDYEHFEEFKKPLYELHAYVRAKLMN 249  
 QY 251 AYPSTISPIGCLPAHLIGDMWGRFTNLVSLTVPFGOKPNIDVTAMDADQRIEKE 310  
 Db 250 AYPSTISPIGCLPAHLIGDMWGRFTNLVSLTVPFGOKPNIDVTAMDADQRIEKE 309  
 QY 311 AEKFEVSVGLPMTQGFENSMETDPGNVOKAVCHPTANLQGGDRILMCTKVTMDDEL 370  
 Db 310 AEKFEVSVGLPMTQGFENSMETDPGNVOKAVCHPTANLQGGDRILMCTKVTMDDEL 369  
 QY 371 TANHENGHIQYDMAYAAQPELLLRNANGCFHEAVGEIMSLSAATPKHLKSGILSPDQGE 430  
 Db 370 TANHENGHIQYDMAYAAQPELLLRNANGCFHEAVGEIMSLSAATPKHLKSGILSPDQGE 429  
 QY 431 DNTEINFLKQALITVGLTPTVYMLERWMTVTEGELPKQOMKMKMEK 481  
 Db 430 DNTEINFLKQALITVGLTPTVYMLERWMTVTEGELPKQOMKMKMEK 480  
 RESULT 12  
 ID AA009102 standard; Protein; 261 AA.  
 XX AA009102;  
 AC AA009102;  
 DT 20-DEC-2001 (first entry)  
 XX  
 DE Novel human protein NHP #11.

XX Human; novel human protein; NHP; antidiabetic; antineumatic;  
 KW antiarthritic; cytostatic; antiarteriosclerotic; vulnerary;  
 KW neuroprotective; nootropic; antiparkinsonian;  
 KW anti-human immunodeficiency virus; antiaesthetic; vasotropic; cardiac;  
 KW hypotensive; anorectic; antilethargy; neuroleptic; anticonvulsant;  
 KW antitumor; immunosuppressive; cerebroprotective; antimicrobial;  
 KW antinflammatory; antibacterial; dermatoprotective; thyromimetic;  
 KW immunomodulator; antiseborrheic; vasodilator; hypertension;  
 KW gastrointestinal disorder; cardiovascular disorder; hyperextension;  
 KW coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;  
 KW cachexia; male infertility; impotence; testicular cancer; lung tumour;  
 KW hyperproliferative disorder; pulmonary system disorder;  
 KW central nervous system disorder; bone disorder;  
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
 KW Huntington's disease; schizophrenia; mania; dementia; paraneoplasia;  
 KW panic disorder; learning disability; immune system disorder;  
 KW psychosia; autism; sleep disorder; immune system disorder;  
 KW Hashimoto's thyroiditis; musculo-skeletal system disorders;  
 KW multiple sclerosis; ischemic brain injury; stroke; infectious disease;  
 KW diabetes mellitus; immunological disorder; asthma; AIDS; immunogen;  
 KW acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;  
 KW inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;  
 KW neural system disorder; respiratory disorder; olfactory disorder;  
 KW wound healing.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2001/74896-A1.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 02-APR-2001; 2001WO-US10542.  
 XX  
 PR 03-APR-2000; 2000US-194118P.  
 XX 29-SEP-2000; 2000US-236384P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;  
 PI Li Y, Dillon PJ;  
 DR WPI: 2001-626394/72.  
 DR N-PDB: AAS14890.  
 XX  
 PT New human proteins, useful for diagnosing, treating, preventing and/or  
 PT prognosing disorders related to the proteins, including cardiovascular  
 PT disorders, autoimmune disorders and reproductive disorders -  
 PS Claim 11; Page 311-312; 318pp; English.  
 XX The invention relates to novel human proteins (NHP) and the  
 CC nucleic acids that encode them and antibodies raised against them.  
 CC The proteins, antibodies and nucleic acids are useful in the diagnosis,  
 CC prognosis, prevention and/or treatment of diseases and/or disorders  
 CC involving vasoconstriction, gastrointestinal disorders, cardiovascular  
 CC disorders (e.g. hypertension, erectile dysfunction, high blood pressure,  
 CC coronary heart disease and arteriosclerosis), anorexia, obesity, bulimia,  
 CC cachexia, disorders of small intestine, disorders of reproductive system  
 CC (e.g. male infertility and/or impotence), testicular cancer, lung tumours  
 CC and other hyperproliferative disorders, disorders of pulmonary system,  
 CC central nervous system disorders, bone disorders, neurodegenerative  
 CC diseases and behavioural disorders (e.g. Alzheimer's disease, Parkinson's  
 CC disease, Huntington's disease, schizophrenia, mania, dementia, paraneoplasia,  
 CC panic disorder, learning disability), immune system disorders (e.g.  
 CC psychosia, autism, sleep disorders), musculo-skeletal system disorders,  
 CC Hashimoto's thyroiditis), renal and musculo-skeletal system disorders,  
 CC central nervous system disorders (e.g. multiple sclerosis, ischaemic  
 CC brain injury and/or stroke), infectious diseases, diabetes mellitus,  
 CC immunological disorders (e.g. asthma, acquired immunodeficient syndrome  
 CC (AIDS), leukaemia, rheumatoid arthritis, inflammatory bowel disease,  
 CC sepsis, acne, psoriasis and lupus erythematosus), neural system  
 CC disorders, respiratory disorders, olfactory disorders and wound



CC healing. The present sequence represents an NHP of the invention.  
 XX  
 SQ Sequence 261 AA;

Query Match 31.7%; Score 1359; DB 22; Length 261;  
 Best Local Similarity 99.6%; Pred. No. 5.3e-107;  
 Matches 252; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 62 MNNAGDKWSAFLEKQSTLAQMPLOEIQNLTVKLOLQALQNGSSVSEDSKRLNTILN 121  
 DB 1 MNNAGDKWSAFLEKQSTLAQMPLOEIQNLTVKLOLQALQNGSSVSEDSKRLNTILN 60  
 QY 122 TMSITSTGKVCNPNDOECLELPEGLNIMANSIDYNERLWAMESRSEYKQRLPYE 181  
 DB 61 TMSITSTGKVCNPNDOECLELPEGLNIMANSIDYNERLWAMESRSEYKQRLPYE 180  
 QY 182 EYVVLKNEARAHNEDYGDYWRGDEYVNGVGDYSGOLIEVEHTEFEIRPLYEHLH 241  
 DB 121 EYVVLKNEARAHNEDYGDYWRGDEYVNGVGDYSGOLIEVEHTEFEIRPLYEHLH 180  
 QY 242 AYVRKALMAYPSYISPIGCLPAHLGDMGREFWNTLSLTVPGQKPNIDVTDAWDOA 301  
 DB 181 AYVRKALMAYPSYISPIGCLPAHLGDMGREFWNTLSLTVPGQKPNIDVTDAWDOA 240  
 QY 302 WDAORIFKEAEKF 314  
 DB 241 WDAORIFKEAEKF 253

## RESULT 13

AA010426  
 ID AAR10426 standard; Protein; 732 AA.  
 XX  
 AC AAR10426;

DT 10-APR-1991 (first entry)  
 XX

DE Human testicular angiotensin conversion enzyme.  
 KW human testicular angiotensin conversion enzyme; TACE;  
 KW male sterility.  
 OS Homo sapiens.

XX  
 XX

Key Location/Qualifiers  
 FT Peptide 1..21  
 FT /label= signal peptide  
 FT Protein 22..732  
 FT /label= mature TACE

PN W09100354-A.  
 XX  
 XX

PD 10-JAN-1991.  
 XX  
 XX

PE 05-JUL-1990; 90WO-FR00513.  
 XX  
 XX

PR 05-JUL-1989; 89FR-0009062.  
 XX  
 XX

PA (INRM ) INST NAT SANTE RECH.  
 XX  
 XX

PI Soubrier F, Alhenc-Gelas F, Hubert C, Corvol P,  
 DR WPT, 1991-036748/05.  
 DR N-PSDB; AAO10328.  
 XX  
 XX

PT Nucleic acid - encoding human testicular angiotensin conversion  
 PT enzyme, used e.g. for in vitro detection of enzyme in organism  
 XX  
 XX

PS Claim 1; Fig 1; 48pp; French.  
 CC  
 CC A bank of human testicular cDNA in Lambda gt11 was screened with a  
 CC probe containing the final 3248 nucleotides of endothelial ACE. The  
 CC complete sequence of TACE was reconstructed from 4 separate clones.

CC The isolated nucleic acid sequence was inserted into a plasmid for  
 CC expression of the protein. The invention covers polypeptides  
 CC containing all or part of TACE sequence. These are useful in  
 CC treatment of inflammation or infectious diseases, especially acute  
 CC pancreatitis, or diseases in which kinins are involved. Antibodies  
 CC against the polypeptides are useful as immunoassay reagents for  
 CC TACE.  
 XX  
 SQ Sequence 732 AA;

Query Match 31.3%; Score 1344; DB 12; Length 732;  
 Best Local Similarity 41.8%; Pred. No. 5e-105;  
 Matches 259; Conservative 119; Mismatches 204; Indels 38; Gaps 10;

QY 15 TAAOS-----TTEOANTFLDKFNEAEDELTYOSSLASNNYNTITEE-----NVQNM 62  
 DB 61 TSAOSPLVLTDEAEASKFVEEDRTSQVWNNRYALAMNNYNTITTSKILQKNQQA 120  
 QY 63 NNAQDKWSAFLEKQSTLAQMPLOEIQNLTVKLOLQALQNGSSVSEDSKRLNTILN 120  
 DB 121 NHT-----LKGTQAKRFVNOLOTTTKRIKKQDLERAPALPAQLEERYKILLD 172  
 QY 123 MSTIYSTGKVCNPNDOECLELPEGLNIMANSIDYNERLWAMESRSEYKQRLPYE 182  
 DB 173 METTYSVATVCHPNG--SCLOLEPDLTVMAATSKRYEDLWAMEGRNDKAGRAIIQFYPK 230  
 QY 183 YVVLKNEARAHNEDYGDYWRGDEYVNGVGDYSGOLIEVEHTEFEIRPLYEHLH 242  
 DB 231 YVELINQAARLNGYVDAGDSKMSYETPSLE-----QDLERLPELOPLILNLHA 280  
 QY 243 AYVRKALMAY-PSYISPIGCLPAHLGDMGREFWNTLSLTVPGQKPNIDVTDAWDOA 301  
 DB 281 YVRKALHNGAOHINLEGPILPAHLGDMGREFWNTLSLTVPGQKPNIDVTDAWDOA 340  
 QY 302 WDAORIFKEAEKFVSGLPNMGFWNSM.LDPEGNOKAVCHPTAWDLGK--DEFRLM 360  
 DB 341 WTPRMKEADDPFTSLGLLPVPEFWMKSMLEKPTDREVVCHASANDFYNGKDFIKQ 400  
 QY 361 CTKYTMDFLTANHEGHIQYDWAYAAOPLFLRANGSPFHEAGELMSAATPKHLKS 420  
 DB 401 CTTVNLEDLVANHEGHIQYFNQYKDLVYALREBANGFHEAGELVALSVSTPKHLHS 460  
 QY 421 IGLSPDFOEDNETEINFLKQALITVGTPTTWLEKRWMEKGEIPIDOMKKWME 480  
 DB 461 LNLSSGSGD--EHDINLAKKALDKIAFIPFSYLVDOMRWRFVDSITKENNNOEMNSI 519  
 QY 481 KREIVGVPEPVPHDETCDPASLEPHVSNDSYFIRYRTLYOFQFOALCOAKHGPILH 540  
 DB 520 RUKYQGLCPVYRTQGDPPGAKFHIPSSVYIRIFYSFIQOFHEALCOAAGHGPILH 579  
 QY 541 KCDISNSTEAGOKLFNMLRLGKSEPMPLALENVYGAKNMVRPLNYPFLPTWLEKDNK 600  
 DB 580 KCDIYQSKKAGQRLATAMKLCFSRKPWEAMOLITGOPNMSASAMLSYKFLDWMRTENE 639  
 QY 601 --NSFVGW--STDMSFYADQS 617  
 DB 640 LHGEKLMPOYMTWTPNSARS 659

## RESULT 14

AA020501  
 ID AAO20501 standard; Protein; 1265 AA.  
 XX  
 AC AAO20501;

DT 27-JUN-2002 (first entry)  
 XX  
 XX

DE Protein of APP related human homologue hcp51674.  
 XX  
 XX

KW Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;  
 KW amyloid precursor protein; tissue-specific expression control; human APP;  
 KW APP pathway modulator; gene therapy.  
 XX





FT Modified-site 1191, 1193  
 FT /label-putative N-glycosylation site  
 FT Modified-site 1225, 1227  
 FT /label-putative N-glycosylation site  
 PN WO9003435-A.  
 XX 05-APR-1990.  
 PD 27-SEP-1989; 89WO-FR00496.  
 XX 27-SEP-1989; 88FR-0012620.  
 PR (INRM) INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE.  
 PA Soudrier F, Alhenc-Gelas F, Hubert C, Corvol P;  
 XX WPI: 1990-132272/17.  
 DR N-PSDB; AA004027.  
 XX  
 PT New DNA encoding human angiotensin converting enzyme used eg in  
 XX diagnosis of hypertension, evaluation of enzyme inhibitors  
 PS Disclosure; ; p; French.  
 CC Human angiotensin converting enzyme hydrolyses angiotensin I and kinins.  
 CC Either intact enzyme or fragments thereof can be used to generate  
 CC antibodies for diagnostic use. Oligonucleotide probes can also be made  
 CC which are complementary to the sequence encoding the enzyme.  
 XX  
 SQ Sequence 1306 AA;  
 Query Match 31.2%; Score 1337; DB 11; Length 1306;  
 Best Local Similarity 41.7%; Pred. No. 4.8e-104;  
 Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;  
 QY 20 TTEBQAKTEFLKFNHAEADLYGSSLASWNTNITEE-----NCONMAGDKMSA 71  
 DB 644 TDEAKSKFVEEDYDTSOVWNEEYAFANMNVNTITETSKILLQNMJAHNT----- 697  
 QY 72 FLKEOSTLAQYPLQETIOLNFKLOALQNGSSVLESEKSKRLTINTSTYSGK 131  
 DB 698 --LKKGTQARKFVNQONTIKRIKKYODLERALPQOELEENKILLDETTYSVAT 755  
 QY 132 VCNPDNPEQELLEPGANETMANSLDYNERLWAMESRSEVGKQRPYEETVYKKNMA 191  
 DB 756 VCHNG--SCLQEPDLTNVMTSRKYEEDLWAMEGRKAGRAILOFPYKVELINDAA 813  
 QY 192 RANHIEDYGDYWRGDEYVNGVDYDSRGQLIEDVHFEETKPLYEHLAAVYRAKLMA 251  
 DB 814 RLNGYVDAGDSWRSKMYETPSLE-----QDLERLFQELQPLYLMLAAVYRRALHRH 863  
 QY 252 Y-PSYISPIGCLPAHLGLDMGREFNTNLYSLTVPGRKPNIDVDAWQMDAQRERKE 310  
 DB 864 YGAQHINLEGPIPAHLGLMMAQOTWSNITYDLYVPPSPASMDTTEAHLKQGWPRRKEKE 923  
 QY 311 AEFKFEVSVGLPMTQGFWENSMULTDGNVOKAVCHPTAMIDGKC-DEPRIMACTKVTMDPF 369  
 DB 924 ADFFETSLILPVPPEFNKSMLEKPTGREGVCHASAMPFYCKDPRKQCTTVNLEDL 983  
 QY 370 LTAHHEMGIQYDAAVAAQPLRLNANGANEHFAVGEIMSLATPRHKLSTGLSPDQ 429  
 DB 984 VVAHHEMGIQYDAAVAAQPLRLNANGANEHFAVGEIMSLATPRHKLSTGLSPDQ 429  
 QY 430 EDNETEINFLKQALTYIGTLPTFYMLEKRWKMYFKEIKDQMKRMKREIYGYVE 489  
 DB 1044 SD-EHDINFELMKMLDKAFIFESYLDVDMRWRFEDGSLTKENYNEHMSRLKATGICLP 1102  
 QY 490 PVPHEETCPDASLFEHVSNDYSFTRYTTLTQFOFQALCOAKKEGSLCDISNSTE 549  
 DB 1103 FVPRQGDGDFGAKFHIPSSPYIRYFVFTIQFQHEALCOAGHTGPHKCDITYOSKE 1162  
 QY 550 AGORLFLNMLRLGKSEPTLALENVGAKMMNAPLNTFEPFLTWLKDONK--NSFVGV 606

DB 1163 AGORLATAMKLGSRPWPAPMOLITGQPNMSASAMLSYFKPLDMLRTNELHGEKLGWP 1222  
 QY 607 STDMSPYADQS 617  
 DB 1223 QYNWTPNSARS 1233  
 RESULT 16  
 AAM68155  
 ID AAM68155 standard; Protein: 1306 AA.  
 AC AAM68155;  
 XX  
 DT 09-NOV-1998 (first entry)  
 XX  
 DE Human angiotensin converting enzyme.  
 XX  
 KW Angiotensin converting enzyme; ACE; hypertension; exercise; human;  
 XX genetic marker.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..29  
 FT Protein 30..1307  
 FT /label= Sig-peptide  
 FT /label= Mat-protein  
 PN WO9831835-A1.  
 XX  
 PD 23-JUL-1998.  
 XX  
 PF 22-DEC-1997; 97WO-US22974.  
 PR 27-MAY-1997; 97US-0048309.  
 PR 16-JAN-1997; 97US-0035382.  
 PA (UYMA-) UNIV MARYLAND BALTIMORE.  
 XX (UYPI-) UNIV PITTSBURGH.  
 PI Ferrell RE, Hagberg JW;  
 DR WPI: 1998-414128/35.  
 DR N-PSDB; AAV41320.  
 XX  
 PT Analysis of genetic markers to identify subjects who will benefit  
 PT from exercise - also assessing risk of cardiovascular disease from  
 PT angiotensin-converting enzyme genotype  
 PS Disclosure; Page 35-41; 61pp; English.  
 XX  
 XX This is human angiotensin converting enzyme (ACE). The ACE gene  
 (see AAV41320) is polymorphic with 2 common alleles (I and D),  
 resulting in 3 genotypes, II, ID and DD. It is an object of the  
 invention to identify individuals possessing a certain genotype and  
 associated ailments, and to determine if the health of that  
 individual can be improved by altering behavior. A claimed method  
 comprises identifying individuals having a certain phenotype,  
 determining the presence or absence of genetic markers associated  
 with the phenotype, and instituting a lifestyle change to exploit  
 or counteract the phenotype expressed by the gene marker. If the  
 phenotype is hypertension, the gene marker is at least one  
 insertion (I) ACE allele and diastolic blood pressure. The gene marker  
 can be identified by PCR amplification (see AAV41321-22) of the  
 appropriate gene region. The general method can be used to  
 identify subjects who will benefit most from physical exercise  
 and also to identify those who are likely to be successful in  
 sports.  
 XX  
 SQ Sequence 1306 AA;







CC In addition to the features in FT, the carboxypeptidase AA sequence  
 CC has 8 potential N-linked glycosylation sites and a potential  
 CC glycosylphosphatidyl inositol anchor sequence similar to that found  
 CC in Bm6. It has significant homology with zinc dependent dipeptidyl  
 CC carboxypeptidases from mammals.

XX  
 SO Sequence 660 AA;

Query Match 24.6%; Score 1057; DB 16; Length 660;  
 Best Local Similarity 37.2%; Pred. No. 1e-80;

Matches 226; Conservative 110; Mismatches 237; Indels 34; Gaps 13;

```

QY 11 LVAVTAQSTIEQAK--TFLDKFNHEADLFYQSSLASWNYNINITEENVQNNNN--AGD 67
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 36 LATLSNVSALEKDEAMGVAFIEGLNDPYTTINNVDSSSSWDYASNTIDYN-QNMSNKYST 94
QY 68 KMSAFLEQSTLAQWYPLQEIQLNLYKQLQALQONGSSVLSSEKSKRLNTLNTMSTIY 127
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 95 EVSKMERQFGITAKRFDMNFKNDLSIKRLEFRVATIGLALPDCKLENAVTSLSKMAIY 154
QY 128 STGKVC--NPDPQECLELLEPGLNEIMANSIDYNERLAWESWRESEVQRLPYEEYV 184
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 155 GSTKVTYVGKDDLP-----LEPDLFRNKKEVGNVYDKLLQTLWLNHNNAVGAIPAKIQTPTI 209
QY 185 VLKNEKARANHYEDYGDYRWGDEYVNGVDYDYSRGQLEDEYEHTEFEIKPLYEHLAYV 244
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 210 KLSNEASASLDGYDNKISAMLSDETE-----NMTETVDKLMEDLSPLYKKLHAYV 259
QY 245 RAKLMNAVPSYISPIGCLPAHLIDGMGRFWNTLY-SLTVPGQKPNIDVTDAWVQAWD 303
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 260 RAKLREIYFGRLEPDTIPAHLLGMMMAQEWGTLXPHLIME--DKP-LDISKTWVEQWD 316
QY 304 AQRIRKAEKFPVSGLEPNTQGFWEENSMLTDPGNVQKAVCHPTAWDLGKG-DEFILMCT 362
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 317 AQKMFHAEDFTSLGIDNMTSEFWKSIILKPED-REIQCHASAMNNGDDFRIMKCT 375
QY 363 KVTMDFTLAHHEMGHIQYDMAAQAQPLIRNGANEGFHEAVGEIMSLAATPKHLKSG 422
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 376 DPSVEELRTVHHEMGHTIEYQYKHLVLLQEGANEGFHEAVGDLIALSVATKTHYKLS 435
QY 423 LLSPEQEDNETEINFLKQALLTVGLPFTYMLEKRWAVFKEIIPKDOIMKRWEMKR 482
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 436 LKRP--TDKYNADVLLMSALDKIAFLPFGYLLDKRWTLITGTFPDKMEKEWEYRI 492
QY 483 EIVGVVEPVPDDETCPPASLHVSNDSYFIRYTRILYQFQFQALCOAK--HEGPL 539
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 493 KYQGVSPVYKKNESFEFGAKYHALVPLRYFAFIIQFQFHEHLCTYAKKVDENHPF 552
QY 540 HKCDISNSTEAGQKLFNMLRIGKSEPTLALENVGAKNNVAPILLNFEPLFTWLKDON 599
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 553 HECDIYGEKNAGDVLKKGSLGSKRPDPVLEIMAGTRQMSASSLKRYEPELEKWLDERI 612
QY 600 KNSFVGV 606
DB   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 613 KNEVGVW 619

```

Search completed: March 13, 2003, 16:56:56  
 Job time : 50 secs

GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: March 13, 2003, 16:56:13 ; Search time 20 Seconds  
(without alignments)  
3869.410 Million cell updates/sec

Title: US-09-978-385-2  
Perfect score: 4291  
Sequence: 1 MSSSWLLSLVAVTAQAOST.....ISKGENNPQNTDDVQTSF 805

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4281	99.8	804	2	T14762
2	1344	31.3	732	1	S05238
3	1337	31.2	1306	1	A31759
4	1334	31.1	732	1	A35655
5	1334	31.1	1312	1	A34171
6	1334	30.6	1193	2	UC2489
7	1310	30.5	1313	1	UC2038
8	1283	29.9	737	1	A34402
9	1058	24.7	1309	1	S35484
10	1039	24.2	611	2	S65472
11	1030	24.0	615	2	A57533
12	1030	24.0	630	2	UC5374
13	642.5	15.0	907	2	T15792
14	157	3.7	532	2	C83696
15	154	3.6	502	2	AF1310
16	147	3.4	502	2	AE1682
17	139.5	3.3	987	2	AI2011
18	139	3.2	608	2	B82938
19	136	3.2	611	2	B82881
20	135	3.1	501	2	D69943
21	125	2.9	627	2	S40048
22	124.5	2.9	987	2	I48373
23	124	2.9	538	2	E72561
24	123.5	2.9	902	2	E90270
25	123	2.9	990	2	S23416
26	122	2.8	642	2	E98000
27	121	2.8	607	2	AB3511
28	121	2.8	1034	2	T30574
29	120	2.8	1339	2	AB4663

30	119.5	2.8	1283	2	S52500	SMH1 protein homol
31	119.5	2.8	3655	2	T38084	TRAP-like protein
32	119.5	2.8	4540	2	T30838	cytoplasmic dynein
33	118.5	2.8	1780	2	T17272	hypothetical prote
34	118	2.7	642	2	G95129	1,4-alpha-glucan b
35	117	2.7	1575	2	G82905	conserved hypothet
36	116	2.7	963	2	C90535	conserved hypothet
37	116	2.7	1642	2	T08880	NMDA receptor-bind
38	116	2.7	1939	2	D97316	probable S-layer p
39	115.5	2.7	611	2	A75573	probable oligoendo
40	115.5	2.7	3433	1	S28381	utrophin - human
41	115	2.7	524	2	B82202	thermostable carbo
42	115	2.7	950	2	A71655	hypothetical prote
43	115	2.7	952	2	T50451	hypothetical colle
44	114.5	2.7	901	2	AB4653	phage infection pr
45	114.5	2.7	1225	1	B64234	hypothetical prote

## ALIGNMENTS

RESULT 1  
T14762  
hypothetical protein DKFZp434A014.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #extl\_change 20-Sep-1999  
C:Accession: T14762  
R:Wandut, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
A:Reference number: Z18181  
A:Accession: T14762  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-804 <NAN>  
A:Cross-references: EMBL:AL110224  
A:Experimental source: adult testis; clone DKFZp434A014  
C:Genetics:  
A:Note: DKFZp434A014.1

Query Match  
Best Local Similarity 99.8%; Score 4281; DB 2; Length 804;  
Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	2	SSSSWLLSLVAVTAQAOSTIEEQATFTDPRKHEHEDELFFOSSLASMYNTNTEENYON	61
DB	1	SSSSWLLSLVAVTAHSTIEQATFTDPRKHEHEDELFFOSSLASMYNTNTEENYON	60
QY	62	MNNAGDKWSAFLEQOSTLAQMYPLQEIOMLFVKLOLQALQOGSSVLSSEDSKRLNTLN	121
DB	61	MNNAGDKWSAFLEQOSTLAQMYPLQEIOMLFVKLOLQALQOGSSVLSSEDSKRLNTLN	120
QY	122	TMSITVSTGKVCNPNPDECLLEPGLNFIANSLSLDYNRRLAMSWRSENGKOLRPLYE	181
DB	121	TMSITVSTGKVCNPNPDECLLEPGLNFIANSLSLDYNRRLAMSWRSENGKOLRPLYE	180
QY	182	EYVVLKEMARAHYEDYGDYRGDYEVNGVDYDSRQGLIEDYHFEETIKPLYEHLH	241
DB	181	EYVVLKEMARAHYEDYGDYRGDYEVNGVDYDSRQGLIEDYHFEETIKPLYEHLH	240
QY	242	AYVRAKIMNAYPSYISPIGCLPAHLGLDMWGRFTWLYLSITVPEGCKRPIVDTDAMVQA	301
DB	241	AYVRAKIMNAYPSYISPIGCLPAHLGLDMWGRFTWLYLSITVPEGCKRPIVDTDAMVQA	300
QY	302	WDAQRIFREAKRFVSYGLPMTQGFENSMILDPGVCKAACHPTAMDLGKDFRILMC	361
DB	301	WDAQRIFREAKRFVSYGLPMTQGFENSMILDPGVCKAACHPTAMDLGKDFRILMC	360
QY	362	TKVTMDDELTAHHEMGHIOYDMAYAAOPFLRLRNCANGFEHANGETMSLSAATPKIKSI	421
DB	361	TKVTMDDELTAHHEMGHIOYDMAYAAOPFLRLRNCANGFEHANGETMSLSAATPKIKSI	420
QY	422	GLSPDFQEDNETELNFIKLQALITVGTLPPTVYLEKRWKRVKGEIPKQMKMKWEMK	481
DB	421	GLSPDFQEDNETELNFIKLQALITVGTLPPTVYLEKRWKRVKGEIPKQMKMKWEMK	480



Db 421 GLLSPDFQEDNTEINFLKQALITVGLTPTFYMLEKRWVWVKEGELPKDQMAKKNEMK 480  
 QY 482 REIVGVVEVPDEYCDPASLFHVSNDYSPFRYTRRLYLPQFOALCOAAKHEGLHK 541  
 Db 481 REIVGVVEVPDEYCDPASLFHVSNDYSPFRYTRRLYLPQFOALCOAAKHEGLHK 540  
 QY 542 CDSINSTAGOKLFNMLRLGKSEPTLALENVGAKNNVPLNTEPEPTMLKDONK 601  
 Db 541 CDSINSTAGOKLFNMLRLGKSEPTLALENVGAKNNVPLNTEPEPTMLKDONK 600  
 QY 602 SEVGMSTWSPYADOSIVRISLKSALDRAKAYENNDNEMTLFSSVAYAMROYFLKVKNO 661  
 Db 601 SEVGMSTWSPYADOSIVRISLKSALDRAKAYENNDNEMTLFSSVAYAMROYFLKVKNO 660  
 QY 662 MLEFEEDVRANLKPRISFNFVYAPKANSDIIPREVEKAIKMSRSRINDAFRLDND 721  
 Db 661 MLEFEEDVRANLKPRISFNFVYAPKANSDIIPREVEKAIKMSRSRINDAFRLDND 720  
 QY 722 LEFLGIOTPLGPPNPQPVSTWLYEGVYGVYVYVILFTGIRDRKKKKARSGENPY 781  
 Db 721 LEFLGIOTPLGPPNPQPVSTWLYEGVYGVYVYVILFTGIRDRKKKKARSGENPY 780  
 QY 782 ASIDISKGNNPGFQNTDDVQTSF 805  
 Db 781 ASIDISKGNNPGFQNTDDVQTSF 804

## RESULT 2

S05238

peptidyl-dipeptidase A (EC 3.4.15.1) precursor, testicular splice form - human  
 N:Alternate names: angiotensin I-converting enzyme (ACE); CD143; dipeptidyl carboxypeptidase  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Sep-1991 #sequence  
 C:Accession: S05238; #revision 02-Jul-1998 #text\_change 18-Jun-1999  
 R:Letton, A.L.; Soubrrier, F.; Allegri, J.; Hubert, C.; Corvol, P.; Alhenc-Gelas, F.  
 FEBS Lett. 252, 99-104, 1989  
 A:Title: The testicular transcript of the angiotensin I-converting enzyme encodes for the  
 A:Reference number: S05238; MUID:89383720; PMID:2547653  
 A:Accession: S05238  
 A:Molecule type: mRNA  
 A:Residues: 1-732 <EHL>  
 A:Cross-references: EMBL:X16295; NID:q28264; PIDN:CAA3362.1; PID:q28265  
 R:Elhers, M.R.W.; Fox, E.A.; Stridom, D.J.; Riordan, J.F.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 7741-7745, 1989  
 A:Title: Molecular cloning of human testicular angiotensin-converting enzyme: the testis  
 A:Reference number: A33979; MUID:90046671; PMID:2554266  
 A:Accession: A33979  
 A:Molecule type: mRNA  
 A:Residues: 1-732 <EHL>  
 A:Cross-references: GB:M26657; NID:9338666; PIDN:AAA60611.1; PID:9338667  
 A:Experimental source: clones R1.2 and T88  
 C:Comment: neither the complete nucleic acid sequence nor the complete translation are shown  
 C:Genetics: For the renal and pulmonary splice form, see PIR:A31759.  
 A:Gene: GDB:DCP1; ACE  
 A:Cross-references: GDB:119840; OMIM:106180  
 A:Map position: 17q23-17q23  
 C:Function:  
 A:Description: catalyzes the hydrolysis of dipeptides from the carboxyl end of polypeptide  
 C:Keywords: mammalian peptidyl-dipeptidase A  
 C:Keywords: alternative splicing; glycoprotein; metalloproteinase; peptidyl-dipeptide hydrolase  
 F:1-21/Domain: signal sequence #status predicted <STG>  
 F:22-732/Product: peptidyl dipeptidase I #status predicted <MAT>  
 F:736-702/Domain: transmembrane #status predicted <TRM>  
 F:103,121,140,186,368,617,651/Binding site: carbohydrate (asn) (covalent) #status predicted  
 F:414,418,434/Binding site: zinc, catalytic (His, His, Glu) #status predicted  
 F:415/Active site: Glu #status predicted

## Query Match

Best Local Similarity 41.8%; Score 1344; DB 1; Length 732;  
 Matches 259; Conservative 119; Mismatches 204; Indels 38; Gaps 10;

QY 15 TAAOS---TTEQAKTFLEKFNHEADELFYQSLASNNYNNITTE-----NVQNM 62

Db 61 TAAQSPLNLTDAEASKVEEEDRTSQQVNMVEYEAANNNTNTTETSKILLQKNQIA 120  
 QY 63 NNAQKWSAFLEKEOSTLAQMTPLQEIQLVYKLLQALQNGSSVLEEDSKRLNTLNT 122  
 Db 121 NHT-----LKVTQARKFEDVQNLQNTTIKRIKKVQDLERALPAQLEENKLLD 172  
 QY 123 MSTISTGKVCNPDQPCQLLEEGLEINANSJDIYERLMAWSEWSEVQKQLRPLYER 182  
 Db 173 METTYSVAIVCHPQG--SCLOLEPDLTNVATSKRYEDLLMAEGMDKKGRAIIQFPK 230  
 QY 183 YVVLNENAKRANHYEDGYWRGDEYVNGVGYDYSRGQLIEDEYEHFEIRKLYEHLA 242  
 Db 231 YVELINQARANGVVDAGDSWRSWYETPSLE-----QLEELFQELQPLYNLHA 280  
 QY 243 YVRAKIMAY-PSYISPIGLCPALLLDMMGRWTNLSTLPFGQKPNIDVTDAMVDA 301  
 Db 281 YVRALIRHYGAQHTINLEGPPIPAHLGNNMAQWMSNITDLYVFPSPASMDTTEAMLKG 340  
 QY 302 WDAQRIKREAEKFEVSVGLPNMTQGEWNSMLTDPGNVQKAVCHPAPMDLGGK-DEFLIM 360  
 Db 341 WTPRMFKKADDEFTSLGLLPVPEFMKNSMLEKPTDGRVYCHASAMDEYKDFRIKQ 400  
 QY 361 CEKTMDDFLAHHEMHIQYDMAVYAAQPLRLNGANEGFHEAVGEIMSLSATPKHLKS 420  
 Db 401 CTIYNLEDLVAHHEMHIQYFMQYKDLVPLREGANPGFHEALIGVLAIVSTPKHLHS 460  
 QY 421 IGLSPDFQEDNTEINFLKQALITVGLTPTFYMLEKRWVWVKEGELPKDQMAKKNEMK 480  
 Db 461 LNTLSBGSGSD-BHDINFLMKMALDKIATIPSYLDQKRWVFGSTTKKEYNDQEWMSL 519  
 QY 481 KREIVGVVEVPDEYCDPASLFHVSNDYSPFRYTRRLYLPQFOALCOAAKHEGLHK 540  
 Db 520 RLKYGQCPPEVPTQGDPEBPGAKFHFPSSVPYIRVYFSIIQFOFHEALCOAGHTGELH 579  
 QY 541 CDSINSTAGOKLFNMLRLGKSEPTLALENVGAKNNVPLNTEPEPTMLKDONK 600  
 Db 580 KCDIYQSKBAGRLATAMKLGFSRPMELMQITQPMASASMLSYFKPLDMLRTENE 639  
 QY 601 --NSFVGM-STWSPYADOS 617  
 Db 640 LHGEKLGWQYVWTPNSARS 659

## RESULT 3

A31759

peptidyl-dipeptidase A (EC 3.4.15.1) precursor, renal and pulmonary splice form - hum  
 N:Alternate names: angiotensin I-converting enzyme (ACE); CD143; dipeptidyl carboxypeptidase  
 C:Species: Homo sapiens (man)  
 C>Date: 07-Jun-1990 #sequence  
 C:Accession: A31759; P00004  
 R:Soubrrier, F.; Alhenc-Gelas, F.; Hubert, C.; Allegri, J.; John, M.; Tregear, G.; C  
 Proc. Natl. Acad. Sci. U.S.A. 85, 9386-9390, 1988  
 A:Title: Two putative active centers in human angiotensin I-converting enzyme reveal  
 A:Reference number: A31759; MUID:89071703; PMID:2849100  
 A:Accession: A31759  
 A:Molecule type: mRNA  
 A:Residues: 1-1306 <SDU>  
 A:Cross-references: GB:004144; NID:q178285; PIDN:AAA51684.1; PID:q178286  
 A:Experimental source: kidney  
 A:Note: parts of this sequence, including the amino end of the mature protein, were d  
 R:Takeuchi, K.; Shimizu, T.; Ohishi, N.; Seyama, Y.; Takaku, F.; Yoksumoto, H.  
 J. Biochem. 106, 442-445, 1989  
 A:Title: Purification of human lung angiotensin-converting enzyme by high-performance  
 A:Reference number: P00004; MUID:90110025; PMID:2556109  
 A:Accession: P00004  
 A:Molecule type: protein  
 A:Residues: 'XX', 32-34, 'E', 36-37, 'X', 39-41, 'R', 43-46 <TAK>  
 A:Experimental source: lung  
 C:Comment: This splice form is found in many tissues, in particular kidney and lung v  
 C:Genetics:  
 A:Gene: GDB:DCP1; ACE  
 A:Cross-references: GDB:119840; OMIM:106180

A: Map position: 17q23-17q23

A: Function:

A: Description: catalyzes the hydrolysis of dipeptides from the carboxyl end of polypeptide A: Note: plays a role in the control of blood pressure by catalyzing the conversion of an C: Superfamily: mammalian peptidyl-dipeptidase A

C: Keywords: alternative splicing; blood pressure control; glycoprotein; kidney; lung; me

F: 1-29/Domain: signal sequence #status predicted <Sto> F: 30-1306/Product: peptidyl dipeptidase I #status predicted <MAT>

F: 160-1276/Domain: transmembrane #status predicted <TM>

F: 38-54,74,111,146,160,318,445,509,523,677,695,714,760,942,1191,1225/Binding site: carbo

F: 330,394/Binding site: zinc (His) #status predicted

F: 988,992,1008/Binding site: zinc, catalytic (His, His, Glu) #status predicted

F: 989/Active site: Glu #status predicted

Query Match 31.2%; Score 1337; DB 1; Length 1306;

Best Local Similarity 41.7%; Pred. No. 6,3e-84;

Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

20 TIEQAKTFLDKFNHAEDELFGSSLSAENVNTNTEEN-----NVQNNNAGDKWSA 71

644 TDEAKSFVEYEDRTQVWNEEAENWNTNITETSKILLQNMQLANH----- 697

72 FLKEQSTLAOMYPLQEIQNTLVKQLQALQNGSSVLSDEKSKRLNTILMTSTIYSTGK 131

698 --LKYGTQARKFDVNOQNTIKRIKKYQDLERAPQGLEEYKILDMETTYSVAT 755

132 VCNPDNPOECILLERGLNEIMANSLDNERLWAMESRSEVGKQRLPYEEYVLKEMMA 191

756 VCHPMG--SCQLQEPDLTNVATSKYEDLLMAMEGMDKRAIILQFYPRYVELINQAA 813

192 RANHEEDYGDYRGDYEYNGVDYSGQLIEDVEHFEERIKPLYEHLAYYRAKLMA 251

814 RINGVYDAGDSKRSMTFSLP-----QDLERLQYQLOPLYLNLHAYYRSLRHRH 863

252 Y-PSYISPIGCLPAHLIDGMGFEWNTLYSLTPPGQKPNIDVDAMQAMDQRIEKE 310

864 YGAQIINIEGPIPAHLILGNMAQWTMSNIDLVPPSPASMDTEAMLKQGTERRMEKE 923

311 AEKFVSVGLPMTOGFENSMILDPGNVQKAVCHPTAMDLSGK--DEFLIMCTVYTMDF 369

924 ADDEFTSGILLFVPEEFNKSMLKPTDGRVYCHASAMDYNGDFRIKQCTTYVNLDEL 983

370 LTAHEMGIQYDMAAAYAOPLLRNGANGFHEAVGEIMSLAATPKILKISIGLSLDPQ 429

984 VVAHEMGIQYFMOYKQLPALRGANRPFHEAIGDVALSVSPKRLHSLNLLSSGG 1043

430 EDNETEINFLKQALITVGLPTFMLEKRWMMVFKEIPDKQMKKWKKEKRELIVGYE 489

1044 SD--EHIDINFLMKALDKIAIFPSYLVDMRWKRVFDGSITKENYQEWMSLRLKXQGLCP 1102

490 PYPHDETCDDPASTLFHVSNDYSFIYRTTYLQFOFOALCOAAKHEGLPKCDISNSTE 549

1103 PVPRTQGDPEGAKFHHSVPIRTFVSFIQFOFHEALCOAAHGTPLHKCDIYQSK 1162

550 AQOKLFNMLRLGKSEPTLALENVGAKNMNVRPLNTPEPLFTMLKQDK--NSFVGM- 606

1163 AGQRLATAMKLGFSRPMWPMQOLITGPMASAMLSYFKPLLDMLRTEHNLGKELGWP 1222

607 STDMSPYADQS 617

1223 OYNMTPNSARS 1233

RESULT 4

A35655

peptidyl-dipeptidase A (EC 3.4.15.1), testis - mouse

N: Alternate names: peptidyl-dipeptidase I, testis

C: Species: Mus musculus (house mouse)

C: Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C: Accession: A35655

R: Howard, T. E.; Shai, S. Y.; Langford, K. G.; Martin, B. M.; Bernstein, K. E.

Mol. Cell. Biol. 10, 4294-4302, 1990

A: Title: Transcription of testicular angiotensin-converting enzyme (ACE) is initiated w

A: Reference number: A35655; MUID: 90318396; PMID: 2164636

A: Accession: A35655

A: Status: preliminary

A: Molecule type: mRNA

A: Residues: 1-732 <HOW>

A: Cross-references: GB: M55333; NID: g191589; PIDN: AAA37149.1; PID: g191590

C: Superfamily: mammalian peptidyl-dipeptidase A

C: Keywords: alternative splicing; peptidyl dipeptide hydrolase; transmembrane protein;

Query Match 31.1%; Score 1334; DB 1; Length 732;

Best Local Similarity 42.6%; Pred. No. 4.2e-84;

Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

20 TIEQAKTFLDKFNHAEDELFGSSLSAENVNTNTEENNVQNNNAGDKWSAFLKEOSTL 79

69 TDEAKSFVEYEDRTQVWNEEAENWNTNITETSKILLKSTSTVSHLTKYGR 128

80 AQMPLQEIQNTLVKQLQALQNGSSVLSDEKSKRLNTILMTSTIYSTGVCNPDNQ 139

129 AKTFVSNFQNSISIKRIKKQLQDLRAVLPKLEEYNOILDMETTYSLSNICYTNG-- 186

140 ECLLPERGLNEIMANSLDNERLWAMESRSEVGKQRLPYEEYVLKEMARAHYEDY 199

187 TCMPLPEDLTNMATSRYKEELLMAMKSNRDKYGRAILPFPRKYVEFSKIKANGYTDA 246

200 GDYWRGDYEYNGVDYSGQLIEDVEHFEERIKPLYEHLAYYRAKLMAVPS- YISP 258

247 GDSKRSIYESDNLE-----QDLERLQYQLOPLYLNLHAYYRSLRHRHYSYINL 296

259 IGCPLAHLIDGMGFEWNTLYSLTPPGQKPNIDVDAMQAMDQRIEKEKFEYSV 318

297 DGPILPAHLILGNMAQWTMSNIDLVAPPSAPNIDAEAMIKQWTRRLKEKDENFETSL 356

319 GLPMTQGFENSMILDPGNVQKAVCHPTAMDLSGK--DEFLIMCTVYTMDFLTAHENG 377

357 GLPVPPEEFNKSMLKPTDGRVYCHASAMDYNGDFRIKQCTTYVNLDELIAHENG 416

378 HIQYDMAAAYAOPLLRNGANGFHEAVGEIMSLAATPKILKISIGLSLDPQDNTEIN 437

417 HIQYFMOYKQLPALRGANRPFHEAIGDVALSVSPKRLHSLNLLSSGGGSGYEDIN 475

438 FLKQALITVGLPTFMLEKRWMMVFKEIPDKQMKKWKKEKRELIVGYEVPDDEY 497

476 FLKMAALDKIAIFPSYLVDMRWKRVFDGSITKENYQEWMSLRLKQGLCPVPSQGD 535

498 CDPASTLFHVSNDYSFIYRTTYLQFOFOALCOAAKHEGLPKCDISNSTEAGOKLFNM 557

536 FDEPSKTHVPRANVYVYFVSFIQFOFHEALCOAAHGTPLHKCDIYQSKELGKLADA 595

558 LRLGKSEPTLALENVGAKNMNVRPLNTPEPLFTMLKQDK--NSFVGM-STDMS 612

596 MKLGSRFWEPMKMLITGPMASAMLSYFKPLLDMLRTEHNLGKELGWP 653

RESULT 5

A34171

peptidyl-dipeptidase A (EC 3.4.15.1) precursor - mouse

N: Alternate names: ACE; angiotensin-converting enzyme; carboxypeptidase; dipeptidyl

C: Species: Mus musculus (house mouse)

C: Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C: Accession: A34171; A29220; A61477

R: Bernstein, K. E.; Martin, B. M.; Edwards, A. S.; Bernstein, E. A.

J. Biol. Chem. 264, 11945-11951, 1989

J. Biol. Chem. 263, 11021-11024, 1988

A: Title: The isolation of angiotensin-converting enzyme cDNA.

A: Reference number: A29220; MUID: 88298730; PMID: 2841312

A: Molecule type: mRNA

A: Residues: 1-1312 <BER>

A: Cross-references: GB: U04947

R: Bernstein, K. E.; Martin, B. M.; Bernstein, E. A.; Linton, J.; Striker, L.; Striker, G.

J. Biol. Chem. 263, 11021-11024, 1988

A: Title: The isolation of angiotensin-converting enzyme cDNA.

A: Reference number: A29220; MUID: 88298730; PMID: 2841312

Query Match	31.18;	Score 1334;	DB 1;	Length 1312;
Best Local Similarity	42.68;	Pred. No. 1e-83;		
Matches 255;	Conservative 112;	Mismatches 213;	Indels 18;	Gaps 7;

QY	20	TIEQAKTLEDENHAEADLFYSSALSWNNTITEENQONMNAADKKSALFKESJTL	79
Db	649	TDEAKMDRVEEEDRAQVLLNEYAEAWQOYNNTITTEGSKILLEKSTESHNTLKYGR	708
QY	80	AQMYPLOEIONTLVKLOALQONSSVLSIDSKSRKNTLITMSTSTYSGKCNPDNQ	139
Db	709	AKTFDVSNFONSSIKRIIKLOINDRAVLPRKELEBYNOILLMETHYLSJNICYNG--	766
QY	140	ECLLEEGENELMANSLDYNEFLANEMESSEVCKOLRPLXEYVVLKNNEMARANYEDY	199
Db	767	TCPLPEBDLTNNMAASRKYEELLMAMKSMRKVRALRPFPPKRYEVSINKIAKNGYDA	826
QY	200	GDYWRGDYEVNGVDGYDSRCQJLEDYHPTPEELRPLYEHLHAYVRKILMNAVS-YISP	258
Db	827	GDWSRSLYESDULF-----QDEKLYQELQPLYLHLHYVRSLSRHHGSEYTL	876
QY	259	IGCLRAHLGLDNGWGFNTNLVSLVPRGQKNIDVTDAVDAQAMDQRIJFKEAKEFVSU	318
Db	877	DGRIPAHLLGNMMAOTWTSNIYDLVAPPSAPNDIPATEMAIKQCMTPRRJFKEADNFISL	936
QY	319	GLPNTOGFWENSMJLTDGCVNOKAVCPRTAMDLSG-DFTILMCTVYHDDFTLHAHENG	377
Db	937	GLLPVPRPEWMSKMLEKPTDGREVYCHPSAMDYFNGDFRIKQCTSVNMEDLYIAHENG	996
QY	378	HIQYDMAAOPFLRLNANGFHEAAGEJMSLSAATPRKLKISGLISPDPOEDNETELN	437
Db	997	HIQYEMQKDLPYTFREGANPGFHEALGDIMALSVSPKHLVSLNLLSTE-GSGYEYDIN	1055
QY	438	FLKCALTLVGTLPPTVYLERKMMWVKEGELPKDOMMKKWEKREIYVVEPRVPHDETY	497
Db	1056	FLMKMALDLKIAFIEFYSJLIDQMRVRVDSITKENYNOEWSJRLKYGLDCLPVPYRSGD	1115
QY	498	CDPAFLFVHSUNDYEFIRYTYTLQPOFOALQOAKHGEPLHKCDISNSTGAKGLEFM	557
Db	1116	FDPGSKFHPANVPYKRFVPSJIIQOFHBALCRAAGHGPJHKCDIYOSKEAGKLADA	1175
QY	558	LRIGSEPWTLTALNNVGAKMNVPLPLNFEPFLFWLMDK--NSPFGW-STDMSP	612
Db	1176	WKIGYSKWPEAMKLITGQPRMSSAMAMNFKYLTWLTETENRRIGETJGWPEYMWK	1233

Query Match 30.6% Score 1312; DB 2; Length 1193;  
Best Local Similarity 40.4%; Pred. No. 2.9e-82;  
Matches 264; Conservative 110; Mismatches 243; Indels 36; Gaps 10;

A:Reference number: J02489; MUID:95110342; PMID:7811282  
A:Accession: J02489  
A:Residue type: mRNA  
A:Residues: 1-1193 <EST>  
A:Cross-references: GB:IL040175; NID:9685168; PIDN:AAA75554.1; PID:9994708  
C:Comment: This enzyme is a zinc dependant dipeptidyl carboxypeptidase that cleaves a  
C:superfamily: mammalian peptidyl-dipeptidase A  
C:keywords: metal binding; peptididipeptide hydrolase; zinc  
F:316,331,914,929/Binding site: zinc, catalytic (Glu, His, Glu, His) #status predicted

QY 22 EHQATFLDKFHEAEEDLFYQSSLASWNYNTNITEENYONNNNGDKSAFLKEOSTIAQ 81  
DB 544 EAQAKFLESEYSTAEEVYWNATTEASWEYNTNITDHKEVMELEKLANSKTTEGMAR 603  
QY 82 MPELOEIQNLVKKLOLQALQONGSSVLESDSKRLNTLNTMTSTYTSTGYKCPNDN--P 138  
DB 604 QEDPSFODEFQETRIILNKLSTVERALPEDELKEYNTLSDMEPTTYSVAKCRNNFNHP 663  
QY 139 QECILLEPLNIMANSIDYNERLMAHESWSEVSKQRLPLEEYVVLKEMARANYED 198  
DB 664 ----LDDPLDILATSDYNEELFEAMKGWMDASGAKIKDKYKRYVELSNKAALNLYTD 718  
QY 199 YGDYWRGDEYVGVGVGYDYSRQGLLEDEYHFEELKPRYLELHLHYVAKILMNAY--PEYIS 257  
DB 719 NCAYRSLSEYPTPE-----EDLERILQLOPLRYLNLHYVRAALYLNKGAEHIS 768  
QY 258 PIGCLPAHLLDGMGRFNTNLXSLTPFGQKPNIDVDAMVDQAMDQRIFEAKRFPVS 317  
DB 769 LKGPLPAHLNMMMAQSNINFEIDLMPEDPATKVDATPRAMQOGWTPKMFEESDRFFTS 828  
QY 318 VGLPMNTQGFENSNMLDPGVNOKAVCHPTAMD--GQDPRFLMCTVYTDDDLTAHNEH 376  
DB 829 IGLIPMPQEFMDKSLKEPRADGREYVCHASAMDYYNKRDKRKOCTVANNDDLLYHNEH 888  
QY 377 GHIOYDMAAQAPELLRNGANEGEFHEAVGEIIMSLSAATPKHLKSIGLISPFQEDNETEI 436  
DB 889 GHVOYFLIDYMOQPSIFRGANGPGEHAIQDVMAIYSTPRKHLISINLD--QVTENEESDI 947  
QY 437 NEELKQALITGLTPETPMLEKMRWVFKGIPRCDQMKKMEKREITGVVPERPHDET 496  
DB 948 NYLMSIALDKALAFPEGLMDQWRKVKFDRKIKEDENQQMMNLRKYQGLCPRPVRSED 1007  
QY 497 YCDPASLEFHSNDYSFTRYRYTRLYQFOFQALCOAKHGEPLKCDISNTEAGOKLFN 556  
DB 1008 DEDGAKFETIPANPYIRFYSPVIOQFOHOLCKAAGHGRLCTDIYKSKAGKLLGD 1067  
QY 557 MLRLGKSEPTLALENVGAKNNANVRPLNFEPLFTWL--KDQNNKSPFGW--SDWSPY 613  
DB 1068 AMKLGFSKPEEAAQOLITGQPNMSAEALMSYFEPMLTWLKNTNGEVLGWMEYSWTPY 1127  
QY 614 ADQSIKRIKLSKSLG-----DKAEWMDNMELYFRSSVAAMQGYFLKYK 659  
DB 1128 AVTEFHATATDTADPLGMSVCTKOATAGAW----VLLATLAVFLTITSIPLGYK 1175

RESULT 7  
JC02038  
peptidyl-dipeptidase A (EC 3.4.15.1) - rat  
N:Alternate names: angiotensin converting enzyme; kinnase II  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: J02038  
R:Koike, G.; Krieger, J.E.; Jacob, H.J.; Mukoyama, M.; Pratt, R.E.; Dzuu, V.J.  
Biochem. Biophys. Res. Commun. 198, 380-386, 1994  
A:Title: Angiotensin converting enzyme and genetic hypertension: Cloning of rat cDNAs  
A:Reference number: J02038; MUID:94121658; PMID:8292044  
A:Accession: J02038  
A:Molecule type: mRNA  
A:Residues: 1-1313 <KOI>

A:Cross-references: GB:U03734; NID:9437289; PIDN:AAA8211.1; PID:9437290  
 A>Note: the authors translated the codon ACC for residue 159 as Tyr  
 C:Comment: This enzyme is a zinc-containing dicarboxy peptidase that cleaves angiotensin II  
 C:Superfamily: mammalian peptidyl-dipeptidase A  
 C:Keywords: alternative splicing; peptidyl-dipeptide hydrolase; transmembrane protein; 24  
 F:393-400,990-998/Region: catalytic status predicted  
 F:1284-1284/Domain: transmembrane status predicted <TM>

Query Match 30.5%; Score 1310; DB 1; Length 1313;  
 Best Local Similarity 42.0%; Pred. No. 4,7e-82;  
 Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;

20 TIEQAKFLDKFNEHADLFYQSSLASWNTNTTEENONMANNAGDKGSAFLKEOSTL 79  
 DB TDEAKANFEVEYDRTAVLNEVEAEANMNTNTTIGSKILLQKNKEVNHLLKGTW 709  
 QY 80 AOMVPLQEIOMLVKLOLQLOQSSVLSSEKSKRLNTLNTSTIYSTGKVCNPNQ 139  
 DB 710 AKTEVSNFQNSTIKRIKKYQNDRAVLPPNELEEVNQLIDMETTVSANVCYTNG-- 767  
 QY 140 ECLLEPELNEIMNSLDYNERLNAMESEKQKOLRLPEEYVLKNEAARAHYEDY 199  
 DB 768 TOLSEPLNTIMATSRKYEELWVWMSKRWKGRALPPKRYVDSNKTAKLNGSDA 827  
 QY 200 GDYWGDEYVNGVDGYDSRQGLIEDVEHTEIEIKPLYLEHAAVYRAKIMAPPS-YISP 258  
 DB 828 GDSMSSYSESDLE-----QDLKLYQELQPLVNLNHAAYRSLHHHGYSEYINL 877  
 QY 259 IGCPLAHLGDMGFEWNTNLSLVFPGQKPNIDVDAVDAQADQRIKFEAKFEVSV 318  
 DB 878 DGPPIAHLGNMMAQOTWSNIYDLVAPFSPASIDATEAMIKQWTPRIKFEADNFTSL 937  
 QY 319 GLPNTOGFENSMLETDGNYOKAVCHPTAMDLGK--DEFLMCTKYVMDPEFLAHNMG 377  
 DB 938 GLPPPEPFWKSMLEKPTDGEVYCHASAMDFYNGKQFRKOCYVMEELVLAHBMG 997  
 QY 378 HIOYMAVAAOPEFLRNGANEGHEAVGEIMSLSAATPKHLSIGLSPDEQENETEN 437  
 DB 998 HIOYMOYKDLVPTREGANFGHEALIGDVALSYTPKHLHSLNLSS--GSGEHNIN 1056  
 QY 438 FLKALITVGLPTFTMLEKRWVFKGEPKIDQMKKMKWEKREIVGVVPPHDETY 497  
 DB 1057 FLMKALIKIAFIPPSYLDIOWRMVFGSITKEVYNQEMWSLRKLYGGLPPVPRSGD 1116  
 QY 498 CDPALEFVSNDSFIRYRTFLYQFOFQALCOAKHEGLHCKDISNTEAGOKLEMM 557  
 DB 1117 FDPGKFEVPAWVPIRIFISITIOFHEALCRAAGHTGPKCDIYOSKEAGLADA 1176  
 QY 558 IRLGKSEPTLALENVGAKNNVAPRLNTEPELFTWLKDNK--NSFVGN-STDWSP 612  
 DB 1177 KLGYSKQWPEAMKITGQPNMSASAINVFKPLTEMLVTENRHRRGETLGPETWTP 1234

RESULT 8  
 A34402  
 N:peptidyl-dipeptidase A (EC 3.4.15.1) Precursor, testicular - rabbit  
 M:Alternative names: angiotensin I-converting enzyme; dipeptidyl carboxypeptidase I; pepti  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A34402; A60724; A36232; C18700  
 J:Kumar, R.S.; Kusari, J.; Roy, S.N.; Soiffer, R.L.; Sen, G.C.  
 J: Biol. Chem. 264, 16754-16758, 1989  
 A:Title: Structure of testicular angiotensin-converting enzyme. A segmental mosaic isozy  
 A:Reference number: A34402; MUID:89380303; PMID:2550457  
 A:Accession: A34402  
 A:Molecule type: mRNA  
 A:Residues: 1-737 <KDD>  
 A:Cross-references: GB:U05041; NID:9164744; PIDN:AAA31153.1; PID:9164745  
 J:Sen, G.C.; Trekkumkara, T.J.; Kumar, R.S.  
 J: Cardiovasc. Pharmacol. 16(Suppl. 4), S14-S18, 1990  
 A:Title: Angiotensin-converting enzyme: structural relationship of the testicular and the  
 A:Reference number: A60724; MUID:91155372; PMID:1705622

A:Accession: A60724  
 A>Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 73-173 <SEN>  
 A>Note: Identical sequences were obtained for mRNAs from lung and testes  
 R:Chen, Y.N.P.; Riordan, J.F.  
 Biochemistry 29, 10493-10498, 1990  
 A:Title: Identification of essential tyrosine and lysine residues in angiotensin con  
 A:Reference number: A36232; MUID:91104959; PMID:2176870  
 A:Accession: A36232  
 A:Molecule type: protein  
 A:Residues: 154-160,236-242 <CHE>  
 R:Iwata, K.; Lai, C.Y.; El-Dorri, H.A.; Soiffer, R.L.  
 Biochem. Biophys. Res. Commun. 107, 1097-1103, 1982  
 A:Title: The NH2- and COOH-terminal sequences of the angiotensin-converting enzyme is  
 A:Reference number: A90107; MUID:83048249; PMID:6291514  
 A:Accession: C18700  
 A:Molecule type: protein  
 A:Residues: 33-35,'SN','SS','PAEL','737 <IWA>  
 A>Note: several of the amino acids in reported are tentative  
 C:Comment: The pulmonary and testicular isoforms of this enzyme differ substantially  
 C:Superfamily: mammalian peptidyl-dipeptidase A  
 C:Keywords: alternative splicing; peptidyl-dipeptide hydrolase; testis; transmembrane

Query Match 29.9%; Score 1283; DB 1; Length 737;  
 Best Local Similarity 40.8%; Pred. No. 1.4e-80;  
 Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

20 TIEQAKFLDKFNEHADLFYQSSLASWNTNTTEENONMANNAGDKGSAFLKEOSTL 69  
 DB TDEAKANFEVEYDRTAVLNEVEAEANMNTNTTIGSKILLQKNMIAHNTLYGHW 134  
 QY 70 SAPLKESTLAOMVPLQEIOMLVKLOLQLOQSSVLSSEKSKRLNTLNTSTIYST 129  
 DB 135 -----ARRDVSNFQNSTIKRIKKYQNDRAVLPPNELEEVNQLIDMETTVSV 184  
 QY 130 GKYCNPNPQECLELLENEIMANSIDYNERLNAMESEKQKOLRLPEEYVVLKNE 189  
 DB 185 ANVCYVGG--SCQLQEDPLNLMATSKRYDELMLWMTSMKDKGRALILEPFRVYFTNK 242  
 QY 190 MARAHYEDYGDYWRGDEYVNGVDGYDSRQGLIEDVEHTEIEIKPLYLEHAAVRAKIM 249  
 DB 243 AARLNGYVDAGDSRSRYETPTLE-----QDLERLFQLOPLVNLNHAAYGRALH 292  
 QY 250 NAY-PSYISIGCPALHLDGMGFEWNTNLSLVFPGQKPNIDVDAVDAQADQRIE 308  
 DB 293 RHGAQHINLEGPPIAHLGNNMAQOTWSNIYDLVAPFSPASIDATEAMIKQWTPRME 352  
 QY 309 KEAEKFEVSVGLPDMTOGFENSMLETDGNYOKAVCHPTAMDLGK--DEFLMCTKYVTD 367  
 DB 353 EADAKFISLGLVPPPEPFWKSMLEKPTDGEVYCHASAMDFYNGKQFRKOCCTVYME 412  
 QY 368 DFLTAHHEMGIQDMAVAAOPEFLRNGANEGHEAVGEIMSLSAATPKHLSIGLSPD 427  
 DB 413 DLVYVHEMGIQDFMOYKDLPAVALREGANFGHEALIGDVALSYTPKHLHSLNLSS-- 472  
 QY 428 FQEDNETEIMLKQALITVGLPTFTMLEKRWVFKGEPKIDQMKKMKWEKREIVGV 487  
 DB 473 -GGYEHDINFLMMAIDKTAIFPSTLVDEMWRVFGSITKEVYNQEMWSLRKLYGGL 531  
 QY 488 VEVPDHETCDPALEFVSNDSFIRYRTFLYQFOFQALCOAKHEGLHCKDISNS 547  
 DB 532 CPVAPRSGQGFDPGAKNHIPSPYIRFYSITIOFHEALCRAAGHTGPKCDIYOS 591  
 QY 548 TEAGQKLEPMLRLGKSEPTLALENVGAKNNVAPRLNTEPELFTWLKDNK--KSPFG 605  
 DB 592 KEAGKRLADAMKLGYSKQWPEAMKITGQPNMSASAINVFKPLTEMLVTENRHRRGETL 651  
 QY 606 W-STDWSPYADOS 617  
 DB 652 WPOYTWTPNSARS 664

## RESULT 9

S53484

peptidyl-dipeptidase A (EC 3.4.15.1) precursor, pulmonary splice form - rabbit

N:Alternate names: angiotensin-converting enzyme, dipeptidyl carboxypeptidase I, kininas

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text, change 10-Sep-1999

C:Accession: S53484; A23455; A18700; A38655; A49726; S17509

R:Thakumkara, T.J.; Livingston III, W.; Kumar, R.S.; Sen, G.C.

Nucleic Acids Res. 20, 683-687, 1992

A:Title: Use of alternative polyadenylation sites for tissue-specific transcription of t

A:Reference number: S53484; MUID:9218960; PMID:1311831

A:Accession: S53484

A:Molecule type: mRNA

A:Residues: 1-1309 &lt;THE&gt;

A:Cross-references: EMBL:X62551

R:Iwata, K.; Blacher, R.; Soffer, R.L.; Lai, C.Y.

Arch. Biochem. Biophys. 227, 188-201, 1983

A:Reference number: A23455; MUID:84051289; PMID:6314908

A:Accession: A23455

A:Molecule type: protein

A:Residues: 34-47, N/49-55 &lt;TWA&gt;

A:Experimental source: lung

R:Iwata, K.; Lai, C.Y.; El-Dorri, H.A.; Soffer, R.L.

Biochem. Biophys. Res. Commun. 107, 1097-1103, 1982

A:Title: The NH2- and COOH-terminal sequences of the angiotensin-converting enzyme isozym

A:Reference number: A90107; MUID:83048249; PMID:6291514

A:Accession: A18700

A:Molecule type: protein

A:Residues: 34-44, 754-755, 'L', 757 &lt;TW2&gt;

R:Kumar, R.S.; Thakumkara, T.J.; Sen, G.C.

J. Biol. Chem. 266, 3854-3862, 1991

A:Title: The mRNAs encoding the two angiotensin-converting isozymes are transcribed from

A:Reference number: A38655; MUID:91139683; PMID:1847388

A:Accession: A38655

A:Molecule type: DNA

A:Residues: 1-88 &lt;KUM&gt;

A:Cross-references: GB:M58579

R:Ramchandran, R.; Sen, G.C.; Misano, K.; Sen, I.

J. Biol. Chem. 269, 2125-2130, 1994

A:Title: Regulated cleavage-secretion of the membrane-bound angiotensin-converting enzym

A:Reference number: A49726; MUID:94124368; PMID:8294466

A:Accession: A49726

A&gt;Status: preliminary

A:Molecule type: protein

A:Residues: 34-95 &lt;KIR&gt;

C:Comment: This enzyme converts angiotensin I to angiotensin II in presence of divalent

ver, the enzyme has been found also in renal tubules and intestinal mucosa.

C:Superfamily: mammalian peptidyl-dipeptidase A

C:Keywords: alternative splicing; blood pressure control; chloride; glycoprotein; intest

F:1-33/Domain: signal sequence #status predicted &lt;SIG&gt;

F:34-1309/Product: peptidyl-dipeptidase A, pulmonary #status experimental &lt;MAT&gt;

F:59/79,150,322,448,512,680,698,717,945,1194/Binding site: carboxylate (Asn) (covalent)

Query Match 29.9%; Score 1283; DB 1; Length 1309;

Best Local Similarity 40.8%; Pred. No. 3, 4e-80;

Matches 250; Conservative 112; Mismatches 213; Indels 38; Gaps 10;

20 TIEQAKTFLDKFNHEADLFYQSSLASWNTNTITE-----NVQNMNN--AGDKW 69

Db 647 TDBAASRFVEEYDRSQAVWMEYAFANNNYNTNTTEASKILLQNMQIAHHTLTLYGNW 706

70 SAELEKOSTIAQMYPIQETIONLTVKIQLOLAQNGSSVLSDEKSKFLNTILMTSTTST 129

Db 707 -----ARRFVSNFQNAATSKRIKKVODLQRAVLPVKELEYNQIILDMETISV 756

Qy 130 GWVCNDPNQECLELLEPGNLEIMANSIDYNEIMAMESRSVQGLPIVEEYVLKNE 189

Db 757 ANVCARDG--SCLQLEPDLTNIMATSKYIDELLMWMTSRDVGVAIIPEPKYVEFNK 814

Qy 190 MARANHEYDGYRWGDEYVNGVDYDSRGOLIDEVHEETEIPLEVEHAARAKM 249

Db 815 AARINQYVADGWSMSTETPLF-----QDLRLQLOPLYNLHAYGRALH 864

Qy 250 NMY-PSYSPICPLPAHLIGDMWGRFTNLYSLVPEGQKPNIDVTADMAQDAQRI 308

Db 865 RHVGAQHINLEGPIDPAHLIGDMWGRFTNLYSLVPEGQKPNIDVTADMAQDAQRI 924

Qy 309 KEAEKPFVSGVPLNTOGFWENSMLTDPGNQAKVCHPTADMLDG--PFRILMCKTYMD 367

Db 925 EERDKFFSLGLLPVPEPEWKNKSMLEKPLDGEVYCHASANDFYGKPFRIKQCTYWE 984

Qy 368 DELTAHHEHGHIOYDMAYAAQPFLLRNGANEGFHVGEIMSLSATPKHLKSLGSLPD 427

Db 985 DLVYVHHEHGHIOYDMAYAAQPFLLRNGANEGFHVGEIMSLSATPKHLKSLGSLPD 1044

Qy 428 EQEDNEETELNLLKQALITVGTLPFTYMLEKRWKVFEGEIPKQDMKMKMKREIYGV 487

Db 1045 -GGGEHNDINFLMKKALDKIAFIPPSVLYDEMRWRFVFGSITKENYQEMWSLRKYGL 1103

Qy 488 VEPVHDEYCDPALSLEFVSNDSYFIRYTRILYFOFQOELCAAKHEGLHCKDISNS 547

Db 1104 CPPAPSGQDFDPGAKFHPSSVPIRYFVSFIQFQHEALCKRAAGTGTGLTCDIYQS 1163

Qy 548 TEAGKLEFNLRLGKSEPTLALENVGAKNNVPLNLEPFLTMDKDN--KNSPVG 605

Db 1164 KKAQRRLADAMKLGSKWPEAKMYITQPNMSASAMNNYKPLMDMLLTENGHGERLG 1223

Qy 606 W-STWSPYADGS 617

Db 1224 WPQYTWTPNSARS 1236

## RESULT 10

S65472

peptidyl-dipeptidase A (EC 3.4.15.1) precursor - horn fly

N:Alternate names: angiotensin I-converting enzyme

C:Species: Haematobia irritans (horn fly)

C:Date: 28-Oct-1996 #sequence, revision 13-Mar-1997 #text, change 22-Jun-1999

C:Accession: S65472; S65431

R:Wijffels, G.; Fitzgerald, C.; Gough, J.; Ridding, G.; Elvin, C.; Kemp, D.; Willadsen

Eur. J. Biochem. 237, 414-423, 1996

A:Title: Cloning and characterisation of angiotensin-converting enzyme from the dipte

A:Reference number: S65431; MUID:96215437; PMID:8647080

A:Accession: S65472

A:Molecule type: DNA

A:Residues: 1-611 &lt;WLD&gt;

A:Cross-references: EMBL:L43965; NID:9908759; PIDN:AAA70427.1; PID:9908760

A:Note: the source is designated as Haematobia irritans exigua

A:Accession: S65431

A:Molecule type: protein

A:Residues: 16, 'P', 20-42, 75-77, 'Q', 79-81, 'X', 83-84, 179-190 &lt;MIN&gt;

A:Note: the source is designated as Haematobia irritans exigua

C:Gene: ACE

C:Superfamily: mammalian peptidyl-dipeptidase A

C:Keywords: glycoprotein; metal binding; peptidyl-dipeptidase hydrolase; zinc

F:1-17/Domain: signal sequence #status predicted &lt;SIG&gt;

F:18-611/Product: peptidyl-dipeptidase A #status predicted &lt;MAT&gt;

F:53,196,531/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 24.7%; Score 1058; DB 2; Length 611;

Best Local Similarity 37.2%; Pred. No. 3, 8e-65;

Matches 223; Conservative 106; Mismatches 246; Indels 22; Gaps 7;

8 LLSVAVTAAGSTTEQAKTFLDKFNHEADLFYQSSLASWNTNTITEENVQNMNAGD 67

Db 8 ILGLAVCHGATKEEIVATVEYLONINKELAKHTNVEETESWAVASNTTDENERLREISA 67  
 QY 68 KWSAFLEKQSTLAQMPLOEIQNTLVKLOALQONGSSVLSDEKSKRLTILNTSTIY 127  
 Db 68 ENAFLEKVAKDIQKFMWRVYGSADYRORFKSLKGYSALEPAEDVAELLEVSAMSNF 127  
 QY 128 STGVCPNDPNQEC-ILLEGLINEIMANSIDYNERLMAESWSEVQKOLRPLYEEVYL 186  
 Db 128 AKVAVCDYKDSKQKADSLIDPEIEITIKSRDPBELAKYVQFYDKAGTPRNSNEKVEL 187  
 QY 187 KENMARANHYEDYGDYWRGDYEVNGVDYDYSRGQIIEV-VEHTFEIEKPLYEHLAYVR 245  
 Db 188 NTKSAKLNNFTDGAELWDEYE-----DAFEDQLEAFEDIEKPLYDQVHGYYR 236  
 QY 246 AKLMNAV-PSYISPIGCLPAHLIDGMGRFNTNLSLTVFPGQKPNIDYTDAMVDQMDA 304  
 Db 237 YRLKFFKFGDEVASTYGPRLHLLGNMMAQWSSITADIVSPPEKPLVDVSDENVAQGYTP 296  
 QY 305 QRIKFEAKFFVSYGLPNTQGFENSMITDPGNVOKAVCHP7AMDIG-KGDFRIILCTK 363  
 Db 297 LKMFQMGDFPQSMGKIKLPQEPFMDKSLTEKPDGROLVCHASAMDFYITDDVRIKQCTR 356  
 QY 364 VTMDFTLAHHEMGHIQYDMAVAAQPELLRNGANGEGHEAVGELMSISATPHLKSIGL 423  
 Db 357 VTQDQFFVHHMGHIQYFLQYOHQPFVYRTGANPGHEHAGVDLSVSTPKHLERYGL 416  
 QY 424 LSPQEDNETEINFLKQALTYGTLPTVYMLEKRWAVKGEIPKQDMKKWMEKRE 483  
 Db 417 LK-NVSDNEARINQLELTALDKIVLFPFAFTMDKYRALKRGGADSEMKCAFWKLEB 475  
 QY 484 IGVVEPYPHDETYCDPASFLEHVSNDYSFIRYTRTYLQFOFQALCOA-----KHG 537  
 Db 476 YSGIEPPVYRREKPDADAKYHVSADYELKIVLSFIQFQYKSACITAEVYPNQREY 535  
 QY 538 PLHKCDISNSTEAGOKLFNMLRLKSEBPWTALLENVYAKMMVAPPLINTEPFLTWK 596  
 Db 536 PLNCDIYGSKEAGKLFENMLSLGASKPMDALEAFNGERTMTGKALAEYEPFLRWLE 594

## RESULT 11

A57533  
 peptidyl-dipeptidase A (EC 3.4.15.1) 67k precursor - fruit fly (*Drosophila melanogaster*)  
 N:Alternate names: angiotensin-converting enzyme  
 C:Species: *Drosophila melanogaster*  
 C>Date: 08-Feb-1996 #sequence\_revision 08-Feb-1996 #text\_change 26-Feb-1998  
 C:Accession: A57533  
 R:Corneil, M.J.; Williams, T.A.; Lamango, N.S.; Coates, D.; Corvol, P.; Soubrier, F.; He  
 J. Biol. Chem. 270, 13613-13619, 1995  
 A:Title: Cloning and expression of an evolutionary conserved single-domain angiotensin c  
 A:Reference number: A57533; MUID:95293950; PMID:7775412  
 A:Accession: A57533  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-615 <COR>  
 A:Cross-references: GB:U25344  
 C:Genetics:  
 A:Gene: FlyBase:Ance  
 A:Cross-references: FlyBase:FBgn0012037  
 C:Superfamily: mammalian peptidyl-dipeptidase A  
 C:Keywords: peptidyl-dipeptide hydrolase

Query Match 24.2%; Score 1039; DB 2; Length 615;  
 Best Local Similarity 35.8%; Pred. No. 7.8e-64;  
 Matches 219; Conservative 120; Mismatches 251; Indels 22; Gaps 9;

QY 8 LLSLVAVTAQSTIEQAKTFIDKFNHEAEDLFYSSLASWYNTNTEENYQNNMNA 67  
 Db 8 LLSLVAVTAQSTIEQAKTFIDKFNHEAEDLFYSSLASWYNTNTEENYQNNMNA 67  
 QY 68 KWSAFLEKQSTLAQMPLOEIQNTLVKLOALQONGSSVLSDEKSKRLTILNTSTIY 127  
 Db 68 KWSAFLEKQSTLAQMPLOEIQNTLVKLOALQONGSSVLSDEKSKRLTILNTSTIY 127  
 QY 68 ELAKFEKVASDTTKFQWRSYQSEDLKROFKALTKLGYALPDEDYAEILLDTLSAESNF 127

QY 128 STGVCPNDPNQEC-ILLEGLINEIMANSIDYNERLMAESWSEVQKOLRPLYEEVYL 186  
 Db 128 AKVAVCDYKDSKQKADSLIDPEIEITIKSRDPBELAKYVQFYDKAGTPRNSNEKVEL 187  
 QY 187 KENMARANHYEDYGDYWRGDYEVNGVDYDYSRGQIIEV-VEHTFEIEKPLYEHLAYVR 246  
 Db 188 NTKSAKLNNFTDGAELWDEYE-----DAFEDQLEAFEDIEKPLYDQVHGYYR 236  
 QY 246 AKLMNAV-PSYISPIGCLPAHLIDGMGRFNTNLSLTVFPGQKPNIDYTDAMVDQMDA 305  
 Db 237 YRLKFFKFGDEVASTYGPRLHLLGNMMAQWSSITADIVSPPEKPLVDVSDENVAQGYTP 296  
 QY 305 QRIKFEAKFFVSYGLPNTQGFENSMITDPGNVOKAVCHP7AMDIG-KGDFRIILCTK 363  
 Db 297 LKMFQMGDFPQSMGKIKLPQEPFMDKSLTEKPDGROLVCHASAMDFYITDDVRIKQCTR 356  
 QY 364 VTMDFTLAHHEMGHIQYDMAVAAQPELLRNGANGEGHEAVGELMSISATPHLKSIGL 423  
 Db 357 VTQDQFFVHHMGHIQYFLQYOHQPFVYRTGANPGHEHAGVDLSVSTPKHLERYGL 416  
 QY 424 LSPQEDNETEINFLKQALTYGTLPTVYMLEKRWAVKGEIPKQDMKKWMEKRE 483  
 Db 417 LK-NVSDNEARINQLELTALDKIVLFPFAFTMDKYRALKRGGADSEMKCAFWKLEB 475  
 QY 484 IGVVEPYPHDETYCDPASFLEHVSNDYSFIRYTRTYLQFOFQALCOA-----KHG 537  
 Db 476 YSGIEPPVYRREKPDADAKYHVSADYELKIVLSFIQFQYKSACITAEVYPNQREY 535  
 QY 538 PLHKCDISNSTEAGOKLFNMLRLKSEBPWTALLENVYAKMMVAPPLINTEPFLTWK 596  
 Db 536 PLNCDIYGSKEAGKLFENMLSLGASKPMDALEAFNGERTMTGKALAEYEPFLRWLE 594

## RESULT 12

JCS374  
 angiotensin-converting enzyme-related protein - fruit fly (*Drosophila melanogaster*)  
 C:Species: *Drosophila melanogaster*  
 C>Date: 28-May-1997 #sequence\_revision 18-Jul-1997 #text\_change 21-Jul-2000  
 C:Accession: JCS374  
 R:Taylor, C.A.M.; Coates, D.; Shitras, A.D.  
 Gene 181, 191-197, 1996  
 A:Title: The Acer gene of *Drosophila* codes for an angiotensin-converting enzyme homo  
 A:Reference number: JCS374; MUID:97128790; PMID:8973330  
 A:Accession: JCS374  
 A:Molecule type: mRNA  
 A:Residues: 1-630 <RAY>  
 A:Cross-references: EMBL:X96913; NID:g1405881; PIDN:CA65632.1; PID:g1405882  
 C:Genetics:  
 A:Gene: Acer  
 C:Superfamily: mammalian peptidyl-dipeptidase A

Query Match 24.0%; Score 1030; DB 2; Length 630;  
 Best Local Similarity 35.6%; Pred. No. 3.4e-63;  
 Matches 219; Conservative 119; Mismatches 260; Indels 24; Gaps 10;

QY 6 WLLSLVAVTAQSTIEQAKTFIDKFNHEAEDLFYSSLASWYNTNTEENYQNNMNA 65  
 Db 16 WLLSLVAVTAQSTIEQAKTFIDKFNHEAEDLFYSSLASWYNTNTEENYQNNMNA 65  
 QY 66 GDKWSAFLEKQSTLAQMPLOEIQNTLVKLOALQONGSSVLSDEKSKRLTILNTSTIY 125  
 Db 66 GDKWSAFLEKQSTLAQMPLOEIQNTLVKLOALQONGSSVLSDEKSKRLTILNTSTIY 125  
 QY 75 YARNAELNKLRAQDIKSSDYQSEDAIRGAHLSLGLSALNADYALQONASISMT 134  
 Db 75 YARNAELNKLRAQDIKSSDYQSEDAIRGAHLSLGLSALNADYALQONASISMT 134  
 QY 126 IYSTGVCPNDPNQEC-ILLEGLINEIMANSIDYNERLMAESWSEVQKOLRPLYEEV 184  
 Db 126 IYSTGVCPNDPNQEC-ILLEGLINEIMANSIDYNERLMAESWSEVQKOLRPLYEEV 184  
 QY 135 NVATATVCSYNTNRSDCSITLPHIOERLSHRDPAELAWYRWHNDSGTPMRONFAEY 194  
 Db 135 NVATATVCSYNTNRSDCSITLPHIOERLSHRDPAELAWYRWHNDSGTPMRONFAEY 194  
 QY 185 VLKNEMARANHYEDYGDYWRGDYEVNGVDYDYSRGQIIEV-VEHTFEIEKPLYEHLAYVR 244  
 Db 185 VLKNEMARANHYEDYGDYWRGDYEVNGVDYDYSRGQIIEV-VEHTFEIEKPLYEHLAYVR 244



```

Db 195 RLFKASQNLNGHRSYADYVQYE-----DPEFER-----QLDATEKOLLPLRQLHGV 244
Qy 245 RAKLMNAY--PSYISPIGCLPAHLGDMGFRFTNLXSLYFPGCKNIDVTAMVDQAM 303
Db 245 RFRROHGPDPVAPESNIPISLLGNMNGOSNNELDLTPPEKPFVYAKAMEKQGYT 304
Qy 304 AORIFKEAEKFPVSVGLPNTGOGFVNSMLTDCGNOKAVCHPTADLCK--GDFRLTMC 362
Db 305 VQRLFELGQDFQSLGKRALPSPFNLVLTTRDD--RQVYCHASAMDEFQDSQVRLKCT 363
Qy 363 KATMDPFLTAHMHGHIQYDMAAOPFLLRNGANGFHEAVGEINLSAATPKHLKSTG 422
Db 364 EVDSHFYVYVHHEIGHIYQYLYEQOPAVYRGAPNGFHEAVQDVYALSVMSAKHLKAG 423
Qy 423 LLSPEODENETEINFLKQALITVGTLPFTYMLEKRMVVRKGEIPKQMKKWKEMK 482
Db 424 LIE--NGRLDEKSKINOLFQKALSKYVLPFGVADVYRAVRNDELDESQNMCGWQMS 482
Qy 483 ELIVGVEVPVPHDETCYDPAFLPHVSNDSYFIRYRTLYQFOFQALCOAKHEGP---- 538
Db 483 EFGVGPVPEFTEKEDPDPAKYHIDADVEYLRFAHIFQFPHKYLCKKAKQYAPNNSR 542
Qy 539 --LHKDISNTPAGOKLENNMLRLKSESWTLALENVGAKAMNVRPLINFEPLFTWLK 596
Db 543 LITDNCDDIGSKAAGRSLSQFLSKGNSRMKEVLEETGETEMDPAALIEFPEPLYOWLK 602
Qy 597 DONKNSFVGSTWSP 612
Db 603 QE--NSRLGVPLGMP 616

```

## RESULT 13

```

115792
Hypothetical protein C42D8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
R:Hallsworth, K.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid C42D8.
A:Reference number: 218405
A:Accession: U15792
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-907 <HAL>
A:Cross-references: EMBL:U15792; NID:G1293844; PID:G1293847; PIDN:AAA98719.1; GSPDB:GN00
A:Experimental source: strain Bristol N2; clone C42D8
A:Gene: CESP:C42D8.5
A:Map position: X
A:Introns: 140/3; 170/3; 194/3; 300/2; 467/3; 551/2; 600/2; 697/3; 774/2; 851/3

```

```

Query Match 15.0%; Score 642.5; DB 2; Length 907;
Best Local Similarity 27.0%; Pred. No. 3.6e-36;
Matches 172; Conservative 136; Mismatches 289; Indels 41; Gaps 19;

```

```

Qy 2 SSSSLLLSLYAVTAOSTIEQAKTFLDKFNHEADLFTQSSLASMNNTNTEENVQ 61
Db 160 SSNTKTDNLDPGSIK--BEKLRSWLAGYEAKIRLREVALSGRPFNDASPSIKLA 217
Qy 62 MNNAQDKWAFLEKQSTLAQWYPLQETONLTVKLOLALQNGSSVSEKSKRLNTILN 121
Db 218 IDEAVNLTFVRSIQAKOFDASVYDEKVMQOLGYSEGGASALAPSFADYSQAQA 277
Qy 122 TMSITISGKVCNPNQECLELPEGLNEIANSIDYNERLAWESRSEVQOLRPLYE 181
Db 278 ALNRDSKSTICDKDVPAPCALQKIDMSIFRNEKSDASRLQHLVSVVTAIAKS--KSYN 336
Qy 182 EYVVLKNEAMARANYEDYGDYWRGDEYVNG--VDCYDYSKROLIDVYHTTEIKPLTEHL 240
Db 337 NIITISSEGAFLNFGANGAMRSAPFMSSSVKHAEP--DLNKQIDKISTIOPEYOL 393
Qy 241 HAYVRAKLMNAY--PSYISPIGCLPAHLGDMGFRFTNLXSLYFPGCKNIDVTAMV 298

```

```

Db 394 HAYMRQLAGTISNPFVSKDGPPIPAHLFGSLDGDMSAHAEQIKPFEEES--ETPEAML 451
Qy 299 D-----QAMDAORIFKEAEKFPVSVGLPNTGOGFVNSMLTDPGVQKAVCHP--TANPL-G 352
Db 452 SAFNTQNTTKKMFATATRTKSAEPPLPSPYMTSSIFARVWS--KMICHPAALDMRA 510
Qy 353 KGFRLTMCYKVTMDPFLTAHMHGHIQYDMAAOPFLLRNGANGFHEAVGEINLSA 412
Db 511 PNDFFVKACAOGLGEPDFQASLSLYQYLYQYLYQYLYQYLYQYLYQYLYQYLYQYLY 570
Qy 413 ATPKHLKSTGLSPDQDNTE--INFLKQALITVGTLPFTYMLEKRMVVRKGEIPK 471
Db 571 TNPFLYSOKLVSEHNDIKDSYIINKLYKESLSTFKLPTTIAADWRYELFGYVPMN 630
Qy 472 QMNMKWMEMKREIVGVEVPVPHDETCYDPAFLPH--VSNDSYFIRYRTLYQFOF 525
Db 631 KLNDRMWELIRKRYEGVNSPQRYNTSND--ALHNSVQVHS--PATRLISTYLKFOI 685
Qy 526 QDALCOAA--KHEGPLHKDISNSTEAGOKLENNMLRLKSESWTLALENVGAKAMNVR 582
Db 686 LKALCQRELFWLSBG-----CILSDT--EKLRTMKLGSSITWKLKLEMISKGELDAQ 739
Qy 583 PLNVFPEPLFTWLKDONK--NSFVGSTWSPYADQSI 618
Db 740 PLLEYEPFLINWLNNTNEIDQVVGWDGEGTPEVEEI 777

```

## RESULT 14

```

C83696
Hypothetical protein BH0371 [Imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MIMD:20512582; PMID:11058132
A:Accession: C83696
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-532 <STO>
A:Cross-references: GB:AP001508; GB:BA000004; NID:G10172890; PIDN:BA804090.1; GSPDB:G
A:Experimental source: strain C-125
A:Gene: BH0371

```

```

Query Match 3.7%; Score 157; DB 2; Length 532;
Best Local Similarity 21.1%; Pred. No. 0.0058;
Matches 118; Conservative 83; Mismatches 213; Indels 144; Gaps 29;

```

```

Qy 22 EBAQKTFDKFNHEADLFTQSSLASMNNTNTE--NVQNNMNAQDKWAFLEKQ 77
Db 3 EODIERPLSQNRVBDLPQVLLNHWATYTGEDMSQHSLSSEYWAHFSDESFEQ 62
Qy 78 -----TLAQWYPLQETONLTVKLOLALQNGSSVSEKSKRLNTILNTMSTLY 127
Db 63 VTRFRKIDSLPLKQRRLQDLDDHMKIKNOFE--EGRQOILSE--KTSIHFYTFPOQY 118
Qy 128 STGKVCNPNQECLELPEGLNEIANSIDYNERLAWESRSEVQOLRPLYEYVYLK 187
Db 119 NSGRVSNNE-----LIDLIRYDDHRRKQAWFA--SKEVGKRTREKDLQLIRKR 166
Qy 188 NEMARANYEDYGDYWRGDEYVNGVDSYSGQILIEDEH---FEELKRYELHLNAVY 244
Db 167 NEVARNLGFTF-----YMSATQELDLQOFPAMFEYIKTSSDAFPMI 211
Qy 245 -----RAKLMNAYPSYISPIGCLPAHLGDMGFRFTNLXSLYFPGCK--PNIDVTDA 296
Db 212 KDEIDERAKVLIKIKDDLRF-----WDYVDFPFGQAPSTEIEND- 250
Qy 297 MVDQAMDQAIRKEAEKFPVSVGLPNTGOGFVNSMLTDCGNVOK--AVCHPTAMDLEGKD 355

```



Db 251 -FDSEFKDODLEQVVSQTFQAMLP--IDDILKRSULYPRKNKNPFQFC--TDMD-RKGD 304  
 QY 356 FRIIMCTKVTMDLFLAHMHMGH-IOYDMAAQAQPELLRGANEGHEAVGEINLSAAT 414  
 Db 305 IRYVLINDQSMYWTALHHEFGNAVYFKFIDSLPFLRL-----FH-----SHTLTT 351  
 QY 415 PKHLKSIGLS--PDQGE-----DNET-----EINELLKQALTYGLTFYMLEKRW 461  
 Db 352 EASALFGRMTKMAEWERYFLGIDRETCERIGRMEMKLRQM-VYST-----RW 400  
 QY 462 MV----FKG---ELPKDQMKMKWEMKREIVGVPEPHETCDPASLFHVS-----N 508  
 Db 401 MLTFEFKSLYEDPDODINALMKVLEIYQMAP--PEDTGSPPDMAAKMHFSLAPVYQ 458  
 QY 509 DY-----SEIRYRT 519  
 Db 459 DYLLGEMASQHLHYIKT 476

## RESULT 15

AF1310  
 Probable thermostable carboxypeptidases homolog lmo186 [imported] - Listeria monocytogenes  
 C:Species: Listeria monocytogenes  
 C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
 C:Accession: AF1310  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluter, T.; Simoes, N.; Tlerrrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AF1310  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-502 <GLA>  
 A:Cross-references: GB:NC\_003210; PIDN:CAC99964.1; PID:g16411339; GSPDB:GN00177  
 A:Experimental source: strain EGD-e  
 C:Genetics:  
 A:Gene: lmo186  
 C:Superfamily: Thermus aquaticus carboxypeptidase Tag

Query Match 3.6%; Score 154; DB 2; Length 502;  
 Best Local Similarity 20.1%; Pred. No. 0.0086;  
 Matches 128; Conservative 101; Mismatches 215; Indels 192; Gaps 35;

QY 20 TIEBOAKTFLDKFNHEADLEFYQSSLASWNYNT-----NITEENVQNMNA 65  
 Db 4 TLEEFLLAYIKKMEALEBAL-----ALVYWDLRTGAPAKMEGRSDVIGVLSSEIFNMQTS 59  
 QY 66 GDKSAFLKEOSTLAQMYPLQEIQNLVYKLOLALQONSSVYSEKSKRL-----N 117  
 Db 60 -EEMAAFLAGIN-----QDKENLS-ETIRKTLLE--SQKTYDLNKKIKPEVAEY 106  
 QY 118 TILMTSTYSGVKYCPNDPQECLELLEPGINEMANSIDYNERLWAMESRSEVGKOLR 177  
 Db 107 KLVAAQEAETWTTAREQN-----DRAAFEPPLTKL----- 136  
 QY 178 PLYEEYVVLKNEMARAHYEDYGDYWRGDIYEVNGVDG--DYSQGLIEDVEHTFEELIKP 235  
 Db 137 -----EMKR-----KFEVEY--GYEENKYDILLDQYEPGVVSYVDSVFEKVR- 177  
 QY 236 LYEHHAHVYRAKLMNAVYSYISPGICLPAHLIGDMGRFTNLXSLVPPGQKPNIDVTD 295  
 Db 178 -DGTMA-IREKIENE-----GVKPDATILN 200  
 QY 296 AMVDQAMDQRIFFKAERFEVYVGLPNNTOGFWENSMILDPGNVQAKVCHPTAMDLGKGD 355  
 Db 201 TKISDA-----KQKEFSIRI-LNKKMGFDF-----EAGRIDERY-HPFATGLMTGD 243  
 QY 356 FRIIMCTKVTMDLFLTA-----HHEMGHIQY-----DMAVAAPPELLRGANGEGHEAVG-- 405

Db 244 VRI--TTRYNENDFKMAVFGTHSGHAIVQONDALVGP--LANGASGIESOSLF 299  
 QY 406 -RIM--SLSAAPFKHLKSIGLSLSPDFO-----ED-----NETEINFLKQALTYGLP- 451  
 Db 300 YELIIGSLAFAKSNVADPQATKPAFDQVKLEDFYRAVNVSESSLINLEADTL--TYPL 357  
 QY 452 ---FTYMLEKRWAVYFKGEIPKQDMKKWEMKREIVGVPEPHETCDPASLFHVS 508  
 Db 358 HIMIRELEK---ALINGELEKDLPKAMGDIYEEYLG1---RPNDNTGVLOD1HWAGG 411  
 QY 509 DYSFIRYTRTL-YQFQGEALCQAKHGPRLHKCDISNSTEAGQLE--NMLRLKSGE 564  
 Db 412 DGYEPFSTALGLMYAQFENQW---QKEIPIDAILISDDYSELKTLTEHVHAKFKTK 467  
 QY 565 PWTALENVYGAKNMNVRLNVEPPLFTWLDQNK 600  
 Db 468 KEPELITDTTG--EGINPTYLDDLKRRAYVYQFVK 502

## RESULT 16

AE1682  
 Probable thermostable carboxypeptidases homolog lln199 [imported] - Listeria innocua  
 C:Species: Listeria innocua  
 C>Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
 C:Accession: AE1682  
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schluter, T.; Simoes, N.; Tlerrrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AE1682  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-502 <GLA>  
 A:Cross-references: GB:AL592022; PIDN:CAC97229.1; PID:g16414500; GSPDB:GN00178  
 A:Experimental source: strain C11p11262  
 C:Genetics:  
 A:Gene: lln199  
 C:Superfamily: Thermus aquaticus carboxypeptidase Tag

Query Match 3.4%; Score 147; DB 2; Length 502;  
 Best Local Similarity 19.4%; Pred. No. 0.026;  
 Matches 124; Conservative 95; Mismatches 223; Indels 196; Gaps 33;

QY 20 TIEBOAKTFLDKFNHEADLEFYQSSLASWNYNT-----NITEENVQNMNA 65  
 Db 4 TLEEFLLAYIKKMEALEBAL-----ALVYWDLRTGAPAKMEGRSDVIGVLSSEIFNMQTS 59  
 QY 66 GDKSAFL-----KQOSTLAQMYPLQEIQNLVYKLOLALQONSSVYSEKSKRLMT 119  
 Db 60 -EEMAAFLAGLNDKEN-----LSEITRKTLLE-ESQKTYDLNKKIKPEVAEY 109  
 QY 120 LNTMTSTYSGVKYCPNDPQECLELLEPGINEMANSIDYNERLWAMESRSEVGKOLRPL 179  
 Db 110 AQQ-ETAWTTARQN-----DRAAFEPPLTKL----- 136  
 QY 180 YEEYVVLKNEMARAHYEDYGDYWRGDIYEVNGVDG--DYSQGLIEDVEHTFEELIKP 237  
 Db 137 -----EMKR-----KFEVEY--GYEENKYDILLDQYEPGVVSYVDSVFEKVR- 177  
 QY 238 EHLHAVYRAKLMNAVYSYISPGICLPAHLIGDMGRFTNLXSLVPPGQKPNIDVTD 297  
 Db 178 DGTMA-IREKIENE-----GVKPDATILN 202  
 QY 298 VDOAMDQRIFFKAERFEVYVGLPNNTOGFWENSMILDPGNVQAKVCHPTAMDLGKGR 357  
 Db 203 ISDA-----KQKEFSIRI-LNKKMGFDF-----EAGRIDERY-HPFATGLMTGDVR 245  
 QY 358 ILMCTKVTMDLFLTA-----HHEMGHIQY-----DMAVAAPPELLRGANGEGHEA----- 403

Db 246 I--TTRYENNDKFAVGTIHGSHAIYQNDPAALVGP--LANGASMGHESQSLFYE 301  
 QY 404 -----VGEINLSAATPRKHLSIGLLSPDQEDNETELNFKALITVGL 450  
 Db 302 IITGSSLAWSKSYADFOAITKPAFDHV-----LEDFTYANVISSSSLRIADTL--TY 355  
 QY 451 P----FTYMLERKRWAVFKEGIPKQDMKKWEMKREIYGVVPEPHDETCDPSLFHV 506  
 Db 356 PLHIMIRYELER---ALINGELEVDLPKAWGDKEYEYIGI---RPNDNTNGVLODIIHA 409  
 QY 507 SUNDYSTRYRTYL-YQFQFOALCOAAKHEBPLKCDISNTEAGQKLF---NNLRIGK 562  
 Db 410 GGDGFEPFSYALGLMVAQFYHQM---OKELPIDATIASDNYTELKWTIKVHTEGK 465  
 QY 563 SEPWTALLENVGAKNMNVRLNFEPLFTWLKDONK 600  
 Db 466 TKKPLEILTDITG-EGINFTYLLDLLEKRYAVYQFNK 502

## RESULT 17

peptide synthetase [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: A12011  
 R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriduchi, Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: A12011  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-987 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BA018013.1; PID:g17135467; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all11647

Query Match 3.3%; Score 139.5; DB 2; Length 987;  
 Best Local Similarity 19.6%; Pred. No. 0.24; Mismatches 208; Indels 191; Gaps 26;  
 Matches 117; Soches 80;

QY 52 TITEENYQNNMAGDKMSAFLEQSTLAQMPLOEIONLTVKLOLQALQNGSSVISED 111  
 Db 106 TILTEHOVSQOLVACVQALPLQYVFLDEGPILGEISLT-----QITAFWQS-LSDD 159  
 QY 112 KSKRLTINTMTSTIYSGKYCNPDN-----POCLLEPGLNETMANSLD 157  
 Db 160 ALELCNSPDDLMLVILYTGSGTGRKGVMLNHRGYMNLTMQNTFSLQPGDRAQRTSFC 219  
 QY 158 YNERLWAMESWSEYKQLRPLYEYVVLKNEMARANYEDYGDYWRDGYVNGVDGYD 217  
 Db 220 FDISWVE--IFPTIMSGAATICPQREVVLNPEEFAR-----WIOEIOINWHEHVP 268  
 QY 218 SRGQILEDVEH---TFEEIKRPLEYHLNAYVRAKLMNAVYSISPGLCPAHILGDMNGRF 274  
 Db 269 LFGEFISALENETWSPQLRLWLMFSGEA-----LPMSEFI-----QR 304  
 QY 275 W-----TNLYSLTVPGOKPNIDVTAMVDQANDAO-----RIKFAEKFEVSV- 318  
 Db 305 WIDRHGLTGLANLYGPT-----EASIDVYCHLITERDEBLTQIPGKALIDWYVKVL 359  
 QY 319 ---GLPNNMVG---FWESM-----LFDPGNVQKAVCHPAMLDGCD--FRILMCTKV 364  
 Db 360 DQGMPPVPGNNGELMLGCVOLALGYLNDPEKTAQAFCPNPFTDI--PDYIYRTGDLVKE 418  
 QY 365 TMDDELTAHNEGHIQYDMAYAAQPFLLRNGANEGFHEAVGEISLSAATPKHLKSGIL 424  
 Db 419 LPDGTIEYH---GRIDHQVKI-----RGFRLEGEIESVLTTH--DYREAAAL 462  
 QY 425 SPDFQEDNETELNFKALITVGLTPTTYMLEKRWAVFKEGIPKQDMKKWEMKREI 484

Db 463 AVDIGEGQKRLV-----ACISGKKIKRPFKEYLEOK--- 494  
 QY 485 VGVVEPVPHDETTCDDPASLFHVSNDYSFIRYTYTTLQFOFOALCOAAKHEBPLKCDI 544  
 Db 495 -----LPH---YMLP-----ORFLMDSL--PKHNKGLDKBRKAL 523  
 QY 545 -----SNSTEAGCKLFLMRLGKSEPTLALENVGAKMNVRLNFEPLFTWL 595  
 Db 524 VTQLTSDSPSSPLRLPLRPGRAQRW-----LVKYEFPYQWL 562

## RESULT 18

B82938

zinc metalloproteinase oligoendopeptidase F U0065 [imported] - Ureaplasma urealyticum  
 C:Species: Ureaplasma urealyticum  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: B82938  
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
 submitted to Genbank, February 2000  
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a  
 A:Reference number: A82870  
 A:Accession: B82938  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-608 <GLA>  
 A:Cross-references: GB:AE002106; GB:AF222894; NID:g6899011; PIDN:AAF30470.1; GSPDB:GN  
 A:Experimental source: serovar 3; biovar 1  
 C:Genetics:  
 A:Gene: pepF-1; U0065  
 A:Genetic code: SGC3

Query Match 3.2%; Score 139; DB 2; Length 608;  
 Best Local Similarity 19.1%; Pred. No. 0.13; Mismatches 226; Indels 240; Gaps 32;  
 Matches 132; Conservative 93;

QY 22 EQQAK-----TFLDKENHEAE-----DLFYQSLASWYNTNITEENYQNNM 63  
 Db 28 EKKAKYIKAPFTFLDSKQNFQWQILIEEFTIVANRF--NYVNNLNINNVYD----- 78  
 QY 64 NAGDKMSAFLEQSTLAQMPLO-----ELQNLTVKLOLQALQNGSSV- 107  
 Db 79 ---PKMSW--SQKLSAFYELEBTALSNVSVLANEAKIKETVLSQNLVYTRKNEYIF 133  
 QY 108 -----LSEPKSRRLNTLTNT---MSTIYST-----GKYCNPDNPEQCL 143  
 Db 134 RYQPHILANNQSKLPFTLIRADGCFSTIRSTYNNDMKTSDAIDAGKKNPILKNEAEAFV 193  
 QY 144 LEPLGLNIMANSIDYNERLWAMESWSEYKQLRPLYEYVVLKNEMARANYEDYGDY 203  
 Db 194 HLKSKDRVLKSA--YLSMRYRAYDSRESITKM---LYYVLSL--NQOAKAKNPDV--IA 246  
 QY 204 RGYEVNAGVGYDYSKQILEDVEHFEETKRLPYEHLNAYVRAKLMNAVYSISPGLCP 263  
 Db 247 KAAFED---DVIDKSLITLLYDQVKLYKDTNEDYKKRVRTYLLKILKV--SKIEP----- 295  
 QY 264 AHLGLDMWGRFTNLVSLTVPGOKPNIDVTAMVDQANDAOIRFEAKKFPVSGLPYM 323  
 Db 296 -----WNGLEPL---ISKTDITPEERK--OMALDISLSIAGDE---YVS-----NI 333  
 QY 324 TQGFWENSMLTDPGNVQKAVCHPAMLDGKGFRLMCTKVTMDDELTAHNEGHIQYDM 383  
 Db 334 KRAFDEKWSWMLQKRGKRGVAGISGCTKISKYYILMANTNSLRDQITVHLEHLSMHS- 392  
 QY 384 AYAAQPFLLRNGANEGFHE---AVGEISLSAATPKHLKSGILSPDFQEDNETELNFK 439  
 Db 393 -----LYSNRQKITYSDYKIFYAIIASIS-----NEVYLNLYL 425  
 QY 440 -----LKQALITVGLTPTTYMLEKRWAVK-----GEIPKQDMKKWME 479  
 Db 426 LEKTKNDELKMLILDEMLISGFATTTRQVITSNFEWMLNLSGAPPTADVVKKEYOK 485  
 QY 480 MKREIIVG--VVEP-----VPHDETTCDDPASLFHVSNDYSFIRYTYTTLQFO 523

Db 486 LELETKPIVEDLNSIYSSTPLRIPH-----FYGNFY-----YKK 526  
 QY 524 QFOALCOAKHHEGLKDISNSTEAGOKLFNMLRLG-----KSEPW 566  
 Db 527 AVGVAAIISGR-----VETKVTGAKGVFPLSSGSKDPLDTIKLGVLTFRQAW 580  
 QY 567 TLALENVGAKNMNVRPLNFEPLFTWKD 597  
 Db 581 QEALEIVK-----LWIKD 593

## RESULT 19

zinc metalloproteinase Oligodeopeptidase F 00521 [imported] - Ureaplasma urealyticum  
 C:Species: Ureaplasma urealyticum  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: D82881  
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
 Submitted to Genbank, February 2000  
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
 A:Reference number: A82870  
 A:Accession: D82881  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-611 <GLA>  
 A:Cross-references: GB:AE002150; GB:AF222894; NID:96899515; PIDN:AAF30934.1; GSPDB:GN001  
 A:Experimental source: serovar 3; biovar 1  
 C:Genetics:  
 A:Gene: pepF-2; 00521  
 A:Genetic code: SGC3

Query Match 3.2%; Score 136; DB 2; Length 611;  
 Best Local Similarity 19.5%; Pred. No. 0.2;

Matches 119; Conservative 100; Mismatches 256; Indels 134; Gaps 27;

QY 49 NYNNNTEENYQNNMNGDKRSALFK-EQSTLAQMPLQELQNTYK-----94  
 Db 69 NYVSNKLOTNL-IDNEMLAOSKIEHDOHRVAKIF--INFENLAIKNDLINSYLNSS 124  
 QY 95 -----LQALQLOQSSVLSSEKSRNLTLNTWTI-----YSTG 130  
 Db 125 LIKQLEFEELMREKHLNQOKVYTAISRFSFGDIFDVLDSMQYQDSINKKQ 184  
 QY 131 KVCNPNDQECLELEPGINEMANSLDYNERLWAMESRSEVQQLRPLLEIYVKNEM 190  
 Db 185 KVCERN-QTDLVYAKSNDRALKRSAYESHFKAIYDLRNTFSKL--LYEY-VKONEL 238  
 QY 191 ARAHHYEDYGDYWGDEYVNGVDYDYSRGLIDVEHTEFEIKRPLEYHLYAVYR-AKL 248  
 Db 239 AKLHFKDY-----ISDAFSPKVDKNF-----INHITOTKKRAKG 275  
 QY 249 MNAVPSYISPTGCLPAHLIGDMGRFWNTNLSLTPVPGQKPNIDVTDM-VDOAMDARI 307  
 Db 276 INRYTKYRT-----LFLKQYQLTKVPEPMCKNLDIDKKNFSESANKL 320  
 QY 308 EKEAEKFSVGLPMTQGFENSMILDPGVQCAVCHPTAMIDGKGRFRLMCTKTYMD 367  
 Db 321 TLEAALLGSEYINVQAFNEQWISMPNNKISGASISNTKGLDKIFILMNYDEYN 380  
 QY 368 DFLFHHMGIHOYDMAAOPFLRLRNGANGFEHVAEIMSLSAATPKHLKISGLSPD 427  
 Db 381 SLTIVLHGLSHVH--TFFA-----NOSQEVNMEYEFYELASITNEILMNHHLK-K 431  
 QY 428 FOEDNETEINFLKQALITVLTPTTYMLEKRMWVFKGELPKDQMKKWMEMK-REIVG 486  
 Db 432 YENDOLMLYLIDEMISGFIATTTQALFSNEFWA-----NEMINQGEFSMNKIVL 484  
 QY 487 VVEPVPHEET-----YCDPASLFHVSNDYSFIRYTRRLYQRFQ-VALCOAK 534  
 Db 485 AYLEINHDYTGKYNKKNISKYDEANALILNIPH---FYTGNYFYVKKYVIGQIC--- 536  
 QY 535 HEGPLHKCDI--SNSTEAGOKLFNMLRLGKS-EPWTLALENVGAKNMNVRPLNFEPL 592

Db 537 --GLINAIIRFNKNAKKEKFCFFKSGSLSP--LEFINILDK-INDENVMEEVNIIIF 591  
 QY 593 -TWKDONK 600  
 Db 592 NSMIDYIK 600

## RESULT 20

carboxypeptidase homolog ypwA - Bacillus subtilis

C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 C:Accession: D69943

R:Kunst, F.; Ogasawara, N.; Mosser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Be  
 C.: Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.;  
 A.: Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
 Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Ga  
 A:Authors: Harwood, C.R.; Hensaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,  
 Koehler, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardin  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma  
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scan  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; S  
 Akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiy  
 T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshid  
 A:Authors: Yoshikawa, H.F.; Zumbelto, E.; Yoshikawa, H.; Dancin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil  
 A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-501 <KUN>

A:Cross-references: GB:Z99115; GB:AL009126; NID:92634478; PIDN:CAB14125.1; PID:926344

A:Experimental source: strain 168

C:Genetics:

A:Gene: ypwA

C:Superfamily: Thermus aquaticus carboxypeptidase Tag

Query Match 3.1%; Score 135; DB 2; Length 501;  
 Best Local Similarity 20.9%; Pred. No. 0.18;

Matches 98; Conservative 65; Mismatches 164; Indels 142; Gaps 26;

QY 180 YEERYVL--KNEA--RAHHYEDY-----DYMGDEYVNGVGYDS 218  
 Db 103 YKRYVILCSAETAMEEAKGSDPSLSPYLQLEFNKRFTTYW-----GYO-- 150  
 QY 219 RGQLEDEHTEFEIKRPLEYHLYAVYRAKLMAVPSYISPTGCLPAHLIGDMGRFWNTL 278  
 Db 151 -----EHPYALDLDE--PGYTVAVLD-----QLFAEL 177  
 QY 279 YSLTVPF-----GQPNIDVTAMVDAMDARIKEAEKFSVGLPMTQGFENSS 331  
 Db 178 KEALIPLVKOVTAAGNKP-----DTSFTKAPKKEKOKELSLYFLQELGYF--- 224  
 QY 332 MLDPGVQCAVCHPTAMIDGKGRFRLMCTKTYMDFLA-----HHMGILOYD--MAY 385  
 Db 225 --DGGMLDITV--HPRATYTLNRGDVNR--TTRIDKDFRTALPGTTHGCHALYEONIDE 278  
 QY 386 AAOPLRLRNGANGFEH-----VGE-----IMSLAATPKHLKISGLSPD- 428  
 Db 279 ALSGTNSDASNGIHSSOSLFYENRIGRNKKHWTYKKGIDASPVGKRDISL--DDEV 336  
 QY 429 QEDNETEINFLKQALITVLTPTTYMLEKRMWVFKGELPKDQMKKWMEMKREIVG 486  
 Db 337 RAINEKSPSIFRVEADELTPYLHIIIRYEIEK---AIFSEVSEDLPSLIMOKQODYLG 393  
 QY 487 VVEPVPHEET-----YCDPASLFHVSNDYSFIRYTRRLYQRFQ-VALCOAK 534  
 Db 394 I---TPQTDAGLIDVHWAGDGFPGFPSTALCYMAAOLKQMLDELEFPDALLERGEF 450  
 QY 540 HKCDISNSTEAGOKLFNMLRLGKS-EPWTLALENVGAKNMNVRPLNFEPL 587

Fri Mar 14 10:00:50 2003

us-09-978-385-2.rpr

Page 12

Db 451 HPIK-----OWLEKVVHIGKRRKKPDIDIKDATG-ELNVRYLIDY 490

Search completed: March 13, 2003, 16:58:59  
Job time : 33 secs

Fri Mar 14 10:00:51 2003

us-09-978-385-2.rsp

Page 1

GenCore version 5.1.4-P5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 16:53:22 ; Search time 15 seconds  
(without alignments)  
2225.896 Million cell updates/sec

Title: US-09-978-385-2  
Perfect score: 4291  
Sequence: 1 MSSSSWLLSLVAVTAOST.....ISKGNNGFQNTDVTSTF 805

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1344	31.3	732	1 ACET_HUMAN	P12866 homo sapien
2	1337	31.2	1306	1 ACET_HUMAN	P12861 homo sapien
3	1334	31.1	732	1 ACET_MOUSE	P22867 mus musculu
4	1334	31.1	1312	1 ACET_MOUSE	P09470 mus musculu
5	1312	30.6	1193	1 ACET_CHICK	Q10751 gallus gall
6	1310	30.5	1313	1 ACET_RABIT	P12822 oryctolagus
7	1283.5	28.9	1310	1 ACET_RABIT	P12822 oryctolagus
8	1283	25.4	615	1 ACET_MOUSE	Q10714 drosophila
9	1090	25.4	615	1 ACET_MOUSE	Q10715 haematobia
10	1058	24.7	611	1 ACET_HAIE	P50848 bacillus su
11	135	3.1	501	1 YPWA_BACU	P39118 bacillus su
12	135	2.9	627	1 GIGB_BACU	P30195 staphylococ
13	133	2.9	986	1 EPIB_STAP	Q52847 bacillus me
14	121	2.8	1034	1 BGAL_BACU	Q12451 saccharomy
15	119.5	2.8	1283	1 OSHI_YEAS	Q10064 schizosacch
16	119.5	2.8	3655	1 YAMB_SCHPO	Q27171 paramecium
17	119.5	2.8	4540	1 DTRC_PARE	Q17693 caenorhabdi
18	118	2.7	663	1 AKR9_HUMAN	Q99996 h a-kinase
19	118	2.7	3911	1 UTR0_HUMAN	P46939 homo sapien
20	115.5	2.7	3433	1 Y511_RICPR	Q92436 rickettsia
21	115	2.7	950	1 PIP_LACLA	P49022 lactococcus
22	114.5	2.7	901	1 Y309_MYCO	P47551 mycoplasma
23	114.5	2.7	1225	1 Y309_MYCO	Q94508 schizosacch
24	114	2.6	1284	1 CMRP_SCHPO	P20929 homo sapien
25	113.5	2.6	6669	1 NEBU_HUMAN	Q01738 phanerocae
26	112.5	2.6	773	1 CDH_PIRCH	Q99928 schizosacch
27	109.5	2.6	1398	1 Y01A_SCHPO	Q97928 saccharomy
28	109	2.5	906	1 Y01A_SCHPO	P30822 saccharomy
29	108.5	2.5	1084	1 XPOL_YEAS	P23556 saccharomy
30	108.5	2.5	2167	1 YCS2_YEAS	P15398 schizosacch
31	107.5	2.5	1689	1 RPA1_SCHPO	P11532 homo sapien
32	107.5	2.5	3685	1 DMD_HUMAN	P26193 porcine rot
33	107	2.5	736	1 VPA_KOTFC	

RESULT 1	ACET_HUMAN	STANDARD:	PRT:	732 AA.
ID	ACET_HUMAN			
AC	P22966:			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Angiotensin-converting enzyme, testis-specific isoform precursor (EC 3.4.15.1) (ACE-T) (dipeptidyl carboxypeptidase I) (kininase II).			
DE	DCPI OR DCP OR ACE.			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9004671; PubMed=2554286;			
RA	Ehlers M.W., Fox E.A., Strydom D.J., Riordan J.F.;			
RT	"Molecular cloning of human testicular angiotensin-converting enzyme;			
RT	the testis isozyme is identical to the C-terminal half of endothelial			
RT	angiotensin-converting enzyme.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 86:7741-7745(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89338720; PubMed=2547653;			
RA	Iatton A.L., Soubrier F., Allegrini J., Hubert C., Corvol P.;			
RT	"The testicular transcript of the angiotensin I-converting enzyme			
RT	encodes for the ancestral, non-duplicated form of the enzyme.";			
RT	FEBS Lett. 252:99-104(1989).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANTS P-32; G-49 AND S-712.			
RX	MEDLINE=99251580; PubMed=10319862;			
RA	Rieder M.J., Taylor S.L., Clark A.G., Nickerson D.A.;			
RT	"Sequence variation in the human angiotensin converting enzyme.";			
RT	Nat. genet. 22:59-62(1999).			
RN	[4]			
RP	ZINC-BINDING;			
RX	MEDLINE=91308093; PubMed=1649623;			
RA	Ehlers M.R., Riordan J.F.;			
RT	"Angiotensin-converting enzyme: zinc- and inhibitor-binding			
RT	stochiometries of the somatic and testis isozymes.";			
RT	Biochemistry 30:7118-7126(1991).			
RT	Biochemistry 30:7118-7126(1991).			
RT	"FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF			
RT	THE C-TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE			
RT	VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN. ALSO ABLE TO INACTIVATE			
RT	BRADYKININ, A POTENT VASODILATOR.			
RT	CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,			
RT	oligopeptide-l-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither			
RT	Asp nor Glu. Converts angiotensin I to angiotensin II.			
RT	CORRELATOR: BINDS 1 ZINC ION.			
RT	SUBCELLULAR LOCATION: Type I membrane protein.			
RT	ALTERNATIVE PRODUCTS: 2 isoforms; testis-specific (shown here) and			
RT	somatic (AC P12821), are produced by alternative splicing.			
RT	TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.			
RT	INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL			

----- BELONGS TO PEPTIDASE FAMILY M2 -----

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).

```

CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M26657; AAA60611.1; -.
DR      EMBL; X16295; CAA34362.1; -.
DR      EMBL; AF118569; AA028561.1; -.
DR      PIR; S05238; S05238.
DR      PIR; A33979; A33979.
DR      MEROPS; M02.004; -.
DR      MIM; 106180; -.
DR      InterPro: IPR001548; Peptidase_M2.
DR      InterPro: IPR000130; Zn_MTPeptide.
DR      Pfam: PF01401; Peptidase_M2_1.
DR      PRINTS; PR00791; PEPTIDPTASEA.
DR      ProDom; PD004184; Peptidase_M2_1.
DR      PROSITE; PS00142; ZINC_PROTEASE; 1.
KW      Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;
KW      Glycoprotein; Transmembrane; Tests; Signal; Alternative splicing;
KW      Polymorphism.
FT

```

FT	DOMAIN	32	732	ANGIOTENSIN-CONVERTING ENZYME, TESTIS-SPECIFIC ISOFORM, EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	685	701	POTENTIAL.
FT	DOMAIN	702	732	CYTOPLASMIC (POTENTIAL).
FT	METAL	414	414	ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	415	415	BY SIMILARITY.
FT	METAL	418	418	ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	442	442	ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	457	457	POTENTIAL.
FT	CARBOHYD	103	103	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	121	121	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	140	140	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	368	368	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	617	617	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	32	32	S -> P. /FTID=VAR_011710.
FT	VARIANT	49	49	S -> G. /FTID=VAR_011711.
FT	VARIANT	712	712	R -> S. /FTID=VAR_011712.
SO	SEQUENCE	732 AA;	83330 MM;	80E0D19CFA642313 CRC64;

Query Match	31.3%	Score 1344;	DB 1;	Length 732;
Best Local Similarity	41.8%	Pred. No. 1.3e-86;		
Matches 229;	Conservative 119;	Mismatches 204;	Indels 38;	Gaps 10;
15 TAOS-----TTFQKKEK				

[illegible][illegible]

```

121 NHT-----LKSTQARKEDVNOLOMTTTRIIKKVQDLERALLPAGBLEEYKILLD 172
123 MSTIYSTGKVCNPDNPQECTII PDCI NETA.....

```

173 METTYSVAVCHPNHG--SCQLERDLINMATSRYKEDLLWAMEGWRDRAGRAITIOEVNR 330

183 YVVLKKNEMARANHVEDYGDYRWGDDYEVANGVDGTDYSRGLIEDVEHTFEERIKPLYEHHA 242

243 YVRAKLMNAY-PSYISPIGCLPAHITGDMEGBETMAY VOT  
ODLERLUFQELQPLYLNHA 280

D6	281	YVRRAALHNRHGAACHINILEGP	IRAHNLHGNMAQOTWISNITDVLVFPFSA	SMTTEAMLEKOG	340	
QY	302	WDAQRIFEXEXEFVSVGRPNNT	QGFHWSMLTDEGNYOKAVCHPTADLKCG	-DPRILM	360	
D6	341	WPRRRFRKRADEDFETSLG	LPLVPBPFMYKSMLEKPTDRE	VYCHASAMDSFNGDPRITQ	400	
QY	361	CKRYIMDFULAHNHEMGILOYDMAA	APFLIRNGANGCFHEAAGELISAA	TBPKILKS	420	
D6	401	CTTVMLDELVYAHNHEMGILOYMOY	KDLPVALREGANBGFHEALGDVLA	TSVPKILHS	460	
QY	421	IGILSPDFOEDETEINFILTKOAL	TYGTLPTFTMLEKRWMMYFKG	ELPRDQMMKMWEX	480	
D6	461	INLISSEGGSD-EHDINILMAAL	DKTALFIPFSTVLYDQMRWRF	PDGSIITRENYNQEWMSL	540	
QY	481	KREIVGVEPVPHEDETQCDPAS	LFIHVSNDYSFIYYRRTLYO	POBOALCOAKHNGCPIL	540	
D6	520	RKATQGLCPRPBRTQGDPE	QAKFNHIPSVPYIRFYFSL	IQOFHEALCOAAGHNGPIL	579	
QY	541	KCDISNSTEAGOKLFNNLR	LKGKSEBPTLALLENVYKAKNNAN	VRPLNYFEPDLFTWLDQNK	600	
D6	580	KCDIYQSKAEGORLATMKR	LGFRRPPEAWOILTTGQ	PNNASASAMLSYFEPDL	WLTRENE	639
QY	601	--NSFVGM-STDMSPYADDS	617			
D6	640	LHGKELGWFPYNTWPN	SARS	659		

## RESULT 2

ACE\_HUMAN  
ID ACE\_HUMAN  
AC P12821;  
STANDARD;  
PRF, 1306 AA.

DT	01-OCT-1989	(Rel. 12, Created)
DT	01-OCT-1989	(Rel. 12, Last sequence update)
DT	15-JUN-2002	(Rel. 41, Last sequence update)

DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)  
DE (ACE) (Dipeptidyl carboxypeptidase I) (Kininase II) (Cm142)  
GN DCP1 OR Renin

08 beta UR DEF UR ACE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata: Cranialia: Vertebrata: Mammalia: Primates: Hominidae: Homo  
09

CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=89071703; PubMed=2849100;  
RA Soubrier F Alhosec-  
RA

RA Tregear G., Aghen-Geas F., Hubert C., Allegrini J., John M.,  
RT "Two putative active centers in human angiotensin I-converting  
pm revealed by X-ray crystallography and site-directed mutagenesis."

revealed by molecular cloning".  
Proc. Natl. Acad. Sci. U.S.A. 85:9386-9390(1988).  
[2]

RA Rieder M.J., Taylor S.L., and others  
 MEDLINE=99251580; pubmed=10319862;  
 RX MEDLINE=99251580; pubmed=10319862;  
 RP SEQUENCE FROM N.A., AND VARIANTS S-261; W-561 AND S-1286.  
 RP SEQUENCE FROM N.A., AND VARIANTS S-261; W-561 AND S-1286.

RT "Sequence variation in the human angiotensin converting enzyme.";  
RL Nat. Genet. 22:59-62(1999).  
RN [3]

RP PARTIAL SEQUENCE OF 30-46.  
RC TISSUE=Lung;  
BY

RA  
MEDLINE=90110025; PubMed=2558109;  
RA  
Takeuchi K., Shimizu T., Ohishi N., Seyama Y., Takaku F.,  
RA  
Yotsumoto H.,

RT "Purification of human lung angiotensin-converting enzyme by high-performance liquid chromatography: properties and N-terminal amino acid sequence".

RL J. Biochem. 106:442-445 (1989).  
RN [4]  
RP ZINC-BINDING

RX MEDLINE=91308093; PubMed=1649623;  
RA Ehlers M.R.; Riordan J.F.;  
DN

RL angiotensin-converting enzyme: zinc- and inhibitor-binding  
RT stoichiometries of the somatic and testis isozymes."?  
K1 Biochemistry 30:7118-7126(1991)

FT	VARIANT	361		N.A.	/FtId=VAR_011708
FT	VARIANT	361		R -> S.	/FtId=VAR_011709.
FT	VARIANT	1286	1286	Q -> E (IN REF. 2).	D -> R (IN REF. 2).
FT	CONFLICT	35	35	O -> E (IN REF. 2).	D -> R (IN REF. 2).
FT	CONFICT	42	42		
SQ	SEQUENCE	1306 AA:	149714 MM:	1B33BCA7301AA6MA	CRC64;
 Query Match					
		31.2%:	Score 1337:	DB 1:	Length 1306;
		Best local similarity	41.7%:	Pred. No. 9.4e-86:	
		Matches 255;	Conservative 118;	Mismatches 204;	Indels 34;
		Gaps			
OY	20	TIEOAKTFIDKRNHAEDELFOSSIASNYNTNTEE-----NWQNANNAGDKWSA	71		
Db	644	TDEEASKFEEDRKSQVVMMEYAFAMNNYNITNTTETSKILLQKNQIANHH-----	697		
OY	72	FLKEOSTLAOMYPDLQNLIVYKLQLOALQONGVSVSEDSKRLNTLNTMSTIYSTGK	131		
Db	698	-LKYGQAARKFPVNOLQNTTIKRILKKYODLERALPAOELEENKLLDMETTYEVAT	755		
OY	132	VCGNDNQOELLIPGNTEFMANSLDYNRILWAMSWMSEVGKQLRPLEYEYVLAKNEMA	191		
Db	756	VCHNG--SCLQEPDLTNMAISRKIEDLLWAMEGRMKGRALLQGYPPVELINDOA	813		
OY	192	RANHYEDGDYWRKDYEVCNGVDGYDYSRGOLLIEDVEHTFFEEIKPLIEHLHAAYRAKLMA	251		
Db	814	RNGYVDAGGSWMRYETPSLE-----QDERLFEOELQPLLYLMHAYRRALHRH	863		
OY	252	Y-PETISPICLPRLHLLGDMWGFRMTNIXSLTPREGOKPNIDVTAMVDAMDARLPKE	310		
Db	864	YGAOHINLEGIPRALHLLGNNAQTWSNITVDLVVPFSAPSMTTTAPALKOGWTPRRMKE	923		
OY	311	AERFFVSGLPEPNNTOGFEMWMSMLTDGPNOYAKAVCPHADTLDGK--DEFILMCTKYTMDF	369		
Db	924	ADDEFTSLGILLPVPERFWNNSMLEKPTDREYVCHASANDFYNGKDFRIKOCTVNEIDL	983		
OY	370	LTAHEHMGHIOTDMAYYAQPPLLRNGANESEFHAVEIMSIASAARPKHLSIGLSPFQ	429		
Db	984	VVAHEHMGHIOTFYAQYKDLPALAREGANPGFHEAIDVALISTVPKHLHSINTLSSBG	1043		
OY	430	EDNEHEIFELFKOALTITVGTLPETYMLMEKRWAVYFGGEIPIKDOQMKWMMKRELIVAYE	489		
Db	1044	SD-EHDITFLKMALDKIAFLPEFYILDQMRNVDPGSTITKENYQNEWMSRLKYQCLP	1102		
OY	490	PVPHDETCDPASLEFHVSNDSYFIRYTRTPLYOFQEQEALCOAAKBEGJLHCIDSINTE	549		
Db	1103	PVPRTQGEDFCAGKFIHPSSVPYTRYRVSEFIIQOFHEHALCOQAAGHTGPLHKCDITYOSKE	1167		
OY	550	AOQRLEFNMLRLKGSPMLTALENVVAGKANMNVRPLIEPELTWMLKDNK--NSHFVG-	606		
Db	1163	AAORLATAMKLGFSPEWDEAMQILITGQPNNSASAMLSFYPLDWLTRTENELHGKLGMP	1222		
OY	607	STWDSVPYADS 617			
Db	1223	QYNWTPNSARS 1233			
 RESULT 3					
ACET_MOUSE		STANDARD;	PRT;	732 AA.	
ID	ACET_MOUSE	P22967;			
DT	01-AUG-1991 (Rel. 19,	Created)			
DT	01-AUG-1991 (Rel. 19,	Last sequence update)			
DT	15-JUN-2002 (Rel. 41,	Last annotation update)			
DE	Angiotensin-converting enzyme,	testis-specific isoform precursor			
DE	(EC 3.4.1.15) (ACE-r)	(Dipeptidyl carboxypeptidase I) (kininase II).			
GN	DCPI OR ACE.				
OC	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria;	Kodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxId=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				



```

Db      187  TCWPLEDDLLNNMATSRTKRYEELLNMAKMSRQXVGCALLIPFPFKYVESSNKIAKINGYTDA 246
QY      200  GDYRKGDYEVNGVDGDIYSRGOLLIEDVEHTEFEIRPLYEHLHAAYVAKLNNAPYS-YISP 258
Db      247  GDSWRSYTESDNE-----ODLEKTYOELOPYLTNLHAAYVRSRLHNGVSEIINL 296
QY      259  IGCIPAHLLCDMGREFEWNLYSLSTLPFGQKRYNDVTDAVNDQADADRITKEKEFFVSV 318
Db      297  DGEIPAHLLGNMMAQWNSITLYLVAFPSPAPRIDATEMKTIKOGTPRIKEDADNEFTSL 356
QY      319  GLPFWTQGEWNSMLTDPGNVQKAVCHPTANDLKGK-DERILMCTKYTMDEFLTAAHENG 377
Db      357  GLLPVPEPFNMKSMLEKPTDGEVREYVCHPSAMDVFYNGDFRIKQCTSVNMMDLYIAHENG 416
QY      378  HIYDMAVAAPOLPLRNGANBGFHVAEIMSLSAATPKHLKSLGLISDPFQDNTEIN 437
Db      417  HIQFMYQYDLVYTRREGANPEFHAIIDIMALSVPSPKHLYSINLISLE-EGSYEYDIN 475
QY      438  FILKQALLTVGLPFTYMLEKRWNVFEGEIPKQDMKKNWEKREIYGVPEYDEY 497
Db      476  FLMKMLDKIATIPSYLIDQMRHNVFEGSITKENTYQEMWSIRLKYGGLCPVPBSQSD 535
QY      498  CDPAISLFHNSNDYSIRYRTYRLVLOFQEOALCOAANKHESPLHKDISINSTAGCKLFLNM 557
Db      536  FPDGSEFHPAPVAVPVRYFVSFIIOFQFHEALCRAAGHTGRLTKCDIYQSKKAGKLDA 595
QY      558  LRLGSEPTTALLVENVGAKNNMNRPLIFEPFLTKDQONK--NSPFGW-SIOWSP 612
Db      596  MKLGYSKPEPMKRLITQPMNMSAMMNYFKPLTEMLVYENRHGETLGMPEYMAP 653

RESULT 4
ACE_MOUSE
ID  ACE_MOUSE      STANDARD;      PRT; 1312 AA.
AC  P09470;
DT  01-MAR-1989 (Rel. 10, Created)
DT  01-FEB-1991 (Rel. 17, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)
DE  (ACE) (Dipeptidyl carboxypeptidase I) (kininase II).
GN  DCP1 OR ACE.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=89308599; PubMed=2545691;
RA  Bernstein K.E., Martin B.M., Edwards A.S., Bernstein E.A.;
RT  "Mouse angiotensin-converting enzyme is a protein composed of two
RT  homologous domains.";
RL  J. Biol. Chem. 264:11945-11951(1989).
[2]
RP  SEQUENCE OF 1-332 FROM N.A., AND PARTIAL SEQUENCE.
RX  MEDLINE=88298730; PubMed=2841312;
RA  Bernstein K., Martin B.M., Bernstein E.A., Linton J., Striker L.,
RA  Striker G.;
RT  "The isolation of angiotensin-converting enzyme cDNA.";
RL  J. Biol. Chem. 263:11021-11024(1988).
CC  -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
CC  THE TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE
CC  VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC  -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC  oligopeptide-1-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC  Asp nor Glu. Converts angiotensin I to angiotensin II.
CC  -1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC  -1- ALTERNATIVE PRODUCTS: 2 isoforms; somatic (shown) and testis-
CC  specific (AC P22677); are produced by alternative splicing.
CC  -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: J04946; AAA37147.1; -  
 CC EMBL: J04947; AAA37148.1; -  
 CC EMBL: J03940; AAA37146.1; -  
 DR PIR: A29220; A29220.  
 DR PIR: A34171; A34171.  
 DR MEROPS: M02.001; -  
 DR MEROPS: M02.004; -  
 DR MGD: MGI-87874; Ace.  
 DR InterPro: IPR001548; Peptidase\_M2.  
 DR InterPro: IPR000130; Zn\_MTPeptide.  
 DR Pfam: PF01401; Peptidase\_M2; 2.  
 DR PRINTS: PR00791; PEPTIDASEA.  
 DR ProDom: PD004184; Peptidase\_M2; 2.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 2.  
 KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;  
 KW Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.  
 FT SIGNAL 1 34  
 FT CHAIN 35 1312  
 FT ISOFORM.  
 FT DOMAIN 35 1264  
 FT TRANSMEM 1265 1281  
 FT DOMAIN 1282 1312  
 FT REPEAT 232 588  
 FT REPEAT 830 1186  
 FT METAL 395 395  
 FT METAL 396 396  
 FT METAL 399 399  
 FT METAL 993 993  
 FT METAL 994 994  
 FT ACET\_SITE 997 997  
 FT METAL 59 79  
 FT CARBOHYD 79 116  
 FT CARBOHYD 116 151  
 FT CARBOHYD 151 165  
 FT CARBOHYD 165 323  
 FT CARBOHYD 323 514  
 FT CARBOHYD 514 682  
 FT CARBOHYD 682 700  
 FT CARBOHYD 700 719  
 FT CARBOHYD 719 765  
 FT CARBOHYD 765 947  
 FT CARBOHYD 947 1196  
 FT VARIANT 568 568  
 FT SEQUENCE 1312 AA; 150947 MW; 9C13BB0529AD3755 CRC64;  
 Query Match 31.1%; Score 1334; DB 1; Length 1312;  
 Best Local Similarity 42.6%; Pred. No. 1.5e-85;  
 Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

DB 877 DGPFAHLGNMVAQOTWSNITDVAFFPSAPMIDATEAKIKOGWPRRIFKADNFFETSL 936  
 QY 319 GUPNMTGFWEMSMILDPENVOKAYCHPTAMGLGK-DEPRIMCTKTYMDELFTAHNEMG 377  
 DB 937 GLTPVPEEFWNSMLEKPPDGRVCHPSAMFYFKDPRIQCTSYNMEDLYVHNHEMG 996  
 QY 378 HIQDYMAVAOPELLNGANGNEGFHEAVGEIMSLSATPKHLKISGLSPDQEDNETEIN 437  
 DB 997 HIQYFMQYKDLPTFREGANPGFHEAIGDIMALSTPHTLSLNLSTE-SSGYEYDIN 1055  
 QY 438 FLIKQALTVGTLPEFTYMLEKRMWYFKGEIPKDOMKKWEMKEKEIVGVPEPVHDEY 497  
 DB 1056 FLMKMALDKIAFTPSYLLIDDMRWRFVDSITKRENDQEMSLRLKYGLCPVPSOGD 1115  
 QY 498 CDPAFLFVNSYSEIRYTRTLTLOFOFQALCOAKHKGPIHKCDISNTEAGOKLFNM 557  
 DB 1116 FDEGSKFHPANPVYRVFSFIQFQFHALCRAAGHTGPIHKCDIYSKRAGKLLADA 1175  
 QY 558 LRLGSEPTWLAENYVGAKNMNVRPLNYFEPLFTWLDKDNK-NSFYGV-STDWSP 612  
 DB 1176 MKLGYSKPWEAKMLITGQPNMSASAMNYFKPLTEWLVTENRHGETLGWEYVWAP 1233

RESULT 5  
 ACE\_CHICK STANDARD; PRT; 1193 AA.  
 ID ACE\_CHICK  
 AC Q10751;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Angiotensin-converting enzyme (EC 3.4.15.1) (Dipeptidyl  
 DE carboxypeptidase I) (kininase II) (Fragment).  
 GN DCPI OR ACE.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA MEDLINE=95110342; PubMed=7811282;  
 RA Esther C.R., Thomas K.E., Bernstein K.E.;  
 RT "Chicken lacks the testis specific isoform of angiotensin converting  
 RT enzyme found in mammals";  
 RL Biochem. Biophys. Res. Commun. 205:1916-1921(1994).  
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF  
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE  
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.  
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,  
 CC oligopeptide-I-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither  
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.  
 CC -1- COFACTOR: BINDS 2 ZINC IONS.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;  
 Glycoprotein; Transmembrane; Repeat.  
 FT NON TER 1 1156 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 1157 1173 POTENTIAL.  
 FT DOMAIN 1174 1193 CYTOPLASMIC (POTENTIAL).  
 FT REPEAT 125 481  
 FT REPEAT 723 1079  
 FT METAL 288 288  
 FT ACT SITE 289 289  
 FT METAL 292 292  
 FT METAL 886 886  
 FT ACT SITE 887 887  
 FT METAL 890 890  
 FT CARBOHYD 11 11  
 FT CARBOHYD 60 60  
 FT CARBOHYD 216 216  
 FT CARBOHYD 407 407  
 FT CARBOHYD 447 447  
 FT CARBOHYD 485 485  
 FT CARBOHYD 513 513  
 FT CARBOHYD 555 555  
 FT CARBOHYD 575 575  
 FT CARBOHYD 658 658  
 FT CARBOHYD 1089 1089  
 SQ SEQUENCE 1193 AA; 137820 MW; 954472A1BEA471C7 CRC64;

Query Match 30.6%; Score 1312; DA 1; Length 1193;  
 Best Local Similarity 40.4%; Pred. No. 4.7e-84;  
 Matches 264; Conservative 110; Mismatches 243; Indels 36; Gaps 10;

QY 22 EDQAKFLDKENHEDLFYQSSLASWNTNITEVQNMNAGDKSAFLKEOSTIAQ 81  
 DB 544 EQAHEFLSEYVNSTAEVNNATTEASWEYNTITDHNKEVLEKMLAMSKHTEYGMAR 603  
 QY 82 MYPLGEIONLYKLOALQONGSSVLSDEKSKRNTLITNTSTYSGKVCNPN--P 138  
 DB 604 QDPDPQDEYTRILINKLSTYERALPEDELKEYNTLSDMETTYSAKVCRENTPEP 663  
 QY 139 QECLELPEGLNIMNSIDYNERLWAMSWSEVQKRLPYEEVYLKNEKARANYED 198  
 DB 664 ----LDLPLDILATSRDYNELLFAWKGMWDASGAKIKDKYKRVLSNKAAYNGTID 718  
 QY 199 YGDVARGYEVNGVGYDYRSKQILIEDVHPEETKPLYLHLAVYRAKLMAY-PSYIS 257  
 DB 719 NAYWRSLYETPTFE-----EDLERLYLOPLYLHLAVYRAKLMAY-PSYIS 768  
 QY 258 PIGCLPAHLGDMGREFWNTSLTFPGQKPNIDYDAMVDQMDAQRKEAKFPYVS 317  
 DB 769 LKGPILAHLLGMMMAOSNIFDLMPDPDAKYDAPFAMKQOGWTPKMPMEESDRFETS 828  
 QY 318 VGLPMNTOGEWNSMLTDPGNVOKAVCHPTAMD- GKGFRLILCTKYTMDFTLAHHEM 376  
 DB 829 LGLIPMPEFMDKSWIEKPADGREVYCHASAMDFYNNKDFIKICTYVNMMDLTVIHEM 888  
 QY 377 GHIOYDAMAYAOPLIRGANGEGHEANGELMSIATPKLKSIGLISPDQEDNETEI 436  
 DB 889 GHVOFLQYMDQPLSFRGANGFHEHAGIDYALSVSTPKHLISINLLD-QVTEHESDI 947  
 QY 437 NEFLKQALTYGTLPFTYMLEKRMWYFKGEIPKDDMKWMEKRELVGVVPEVDET 496  
 DB 948 NYLMSIALDKIAFLPFGILMDQWKKVDFGRIKEDENQOQWMLNRLYYOGICLPPVPSSED 1007  
 QY 497 YCDPASLHVSNDSYFIRYTRILYOFQFQALCOAKHEGFLKCDISNSTENGOKLFFN 556  
 DB 1008 DEDGAKFHIPANPYIRYEVSVIYOFQFQALCKAAGHGPHLTCIYLSKEGKGLGD 1067  
 QY 557 MLRGKSPWTLLENVVGAKNNVRLNLYFEPFLTWL--KQONKSEYGVG-STDMSPY 613  
 DB 1068 AMKLGSKPEWEMOLLITGQPNMSAEALMSYFEPFLMTLVKKNTEGNEVGVPEYSWTPY 1127  
 QY 614 ADOSIKVRLISLALG-----DKAYEMNDNEMYLFRSSVAVAMQRYELKVK 659

DB 1128 ATVERHAATDADFLGMSVGTQKATAGAM-----VLALALVFLTITSIFLGK 1175  
 RESULT 6  
 ACE\_RAT  
 ID ACE\_RAT STANDARD: PRT: 1313 AA.  
 AC P47820.  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Angiotensin-converting enzyme, somatic isoform precursor (EC 3.4.15.1)  
 DE (ACE) (dipeptidyl carboxypeptidase I) (kininase II).  
 GN DCP1 OR ACE.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=94121658; PubMed=8292044;  
 RA Kolke G., Krieger J.E., Jacob H.J., Mukoyama M., Pratt R.E.,  
 RA Dzau V.J.;  
 RT "Angiotensin converting enzyme and genetic hypertension: cloning of  
 RT rat cDNAs and characterization of the enzyme."  
 RL Biochem. Biophys. Res. Commun. 198;380-386(1994).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LEW/N; TISSUE=Lung;  
 RA Jafarian-Tehrani M., Listwak S., Barrientos R.M., Michaud A.,  
 RA Corvol P., Sternberg E.M.;  
 RT "Characterization of a missense mutation in the angiotensin  
 RT I-converting enzyme cDNA in exudative inflammation resistant F344/N  
 RT rats."  
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.  
 RL  
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF  
 CC THE TERMINAL HIS-LEU. THIS RESULTS IN AN INCREASE OF THE  
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.  
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,  
 CC oligopeptide-I-xaa-Xbb, when xaa is not pro, and Xbb is neither  
 CC asp nor glu. Converts angiotensin I to angiotensin II.  
 CC -1- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: somatic (shown here) and testis-  
 CC specific; are produced by alternative splicing.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U03708; AAA82110.1;  
 CC EMBL: U03734; AAA82111.1;  
 CC EMBL: AF201332; AAG35597.1;  
 CC  
 CC MEROPS: M02.001;  
 CC  
 CC InterPro: IPR001548; Peptidase\_M2.  
 CC InterPro: IPR000130; Zn\_Mrpeptidase.  
 CC Pfam: PF01401; Peptidase\_M2; 2.  
 CC PRINTS: PR00791; PEPDIPYASRA.  
 CC ProDom: PD004184; Peptidase\_M2; 2.  
 CC PROSITE: PS00142; ZINC\_PROTEASE; 2.  
 CC Hydrolase; Metalloprotease; Carboxypeptidase; Zinc; Dipeptidase;  
 CC Glycoprotein; Transmembrane; Repeat; Signal; Alternative splicing.  
 CC SIGNAL 1 35  
 CC CHAIN 36 1313  
 CC FT  
 CC FT  
 CC FT  
 CC DOMAIN 36 1265  
 CC TRANSMEM 1266 1282  
 CC  
 CC EXTRACELLULAR (POTENTIAL).  
 CC POTENTIAL.

```

FT DOMAIN 1283 1313 CYTOPLASMIC (POTENTIAL).
FT REPEAT 233 589
FT REPEAT 831 1187
FT METAL 396 396
FT ACT_SITE 397 397
FT METAL 400 400
FT METAL 994 994
FT ACT_SITE 995 995
FT METAL 998 998
FT CARBOHYD 44 44
FT CARBOHYD 60 60
FT CARBOHYD 80 80
FT CARBOHYD 117 117
FT CARBOHYD 152 152
FT CARBOHYD 166 166
FT CARBOHYD 324 324
FT CARBOHYD 515 515
FT CARBOHYD 683 683
FT CARBOHYD 701 701
FT CARBOHYD 720 720
FT CARBOHYD 766 766
FT CARBOHYD 948 948
FT CARBOHYD 1197 1197
FT VARIANT 207 207
SQ SEQUENCE 1313 AA; 150907 MW; 80B5D0015F129591 CRC64;

Query Match 30.5%; Score 1310; DB 1; Length 1313;
Best Local Similarity 42.0%; Pred. No. 7, 4e-84;
Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;

QY 20 TIEQAKTELDKNEHAEFLFYQSSLASWNTNTTEENVJNNNNAQKWSAFLEKQSTL 79
D 650 TDEAKANRFEEDRAKAVLMNEYEAAMNHYNNITTESKILLQNKKEVSNTTLKYGW 709
QY 80 AQMYPLQIEONLFVKQLQALQNGSSVYSEDKSKRLNTLTMTSTYTGKVCNDNQ 139
D 710 AKTEFVSNTQNTIKRIKKVQVADVAVLPNPLEEYNOILDMETTYAVANCYNG-- 767
QY 140 ECLLEPGLEIMANSIDYNERLAMESRSEYKQRLRYEERYVYLKMKMARANEY 199
D 768 TCSLEPDLNIMATSKREBELLMWKSMDKGRALFFPKYVDFSKIALNLGYSA 827
QY 200 GDYWRGDYEVNGVDYDYSRGQILIEDEHTEFEIKPLRYEHLNAYRAKLNNAPS-YISP 258
D 828 GDSMRSSYESDLE-----QOLEKLYQELQPLVILNLAHYVRSYLRHNGSEYIML 877
QY 259 IGLPRAHLIGDMGRFWTNLSLTVEFGQKPNIDVDAMVDAQDAQRIJFEKAEKFFSV 318
D 878 DGEIPRALHGNMMAQYWSNITDYLVAFPPSAPSIDATEAMIKQGWTPRRIFKEADNFTSL 937
QY 319 GLPNMTQGEFENSMITDPGNVOKAVCHPTAMDIGK-DPRILMCTKVTMDDPFLTAHENG 377
D 938 GLPVPPEPFENKSMLEKFPIDGRVYVCHASMDPYNKDRIKOCISVNMEEVIAHHENG 997
QY 378 HIQYDVAVAQPLRLNGANEGHEAVGELMSISAATPKHLISIGLSPDFQEDNTEIN 437
D 998 HIQYFYQKDLPTVFREGANPGFHEAIGVIALVSTPKHLISLNLSE-SSGYEHDIN 1056
QY 438 FLIKQALITIGLPEYMLKEMRWVFKGEIPKDOMKMKWMEKRIYGVVEVPYDEY 497
D 1057 FLIKMALDKIAIFPESYLLIDQWRWVFDSTIKENNOEWSIRLKYQIGLCIPVPSQSD 1116
QY 498 CDPAFLFHSNDYSFTRYTYRTLYQFOEALCOAKHSGPLKHCISNSTEAGQKLFPM 557
D 1117 FDEGSKFHPANVPYIRYFISFIQOFHEALCRAAGHGPYLKCDITYSKENGKILADA 1176
QY 558 LRIKSEDEPTALENVVAGKNNANRPLNFEPLFTWLKDQNK--NSFGW-STDWP 612
D 1177 MKLGSKQWPEAMKITGQPNMSASAIMNYFKPLTEMLVTENRHHGETIGWPEYWTWP 1234

```

```

ID ACE_RABIT STANDARD; PRT; 1310 AA.
AC P12822; C02852;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Angiotensin-converting enzyme, somatic isoform precursor (BC 3.4.15.1)
DE (ACE) (Dipeptidyl carboxypeptidase I) (Kininase II).
GN DCPI OR ACE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA MEDLINE=92178960; PubMed=1311831;
RT "Thekkumkara T.J., Livingston W. III, Kumar R.S., Sen G.C.;
RT "Use of alternative polyadenylation sites for tissue-specific
RT transcription of two angiotensin-converting enzyme mRNAs.";
RL Nucleic Acids Res. 20:683-687(1992).
RN [2]
RP REVISIONS.
RA Sen G.C.;
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-88 FROM N.A.
RC TISSUE=Liver;
RA MEDLINE=91139683; PubMed=1847388;
RT Kumar R.S., Thekkumkara T.J., Sen G.C.;
RT "The mRNAs encoding the two angiotensin-converting isozymes are
RT transcribed from the same gene by a tissue-specific choice of
RT alternative transcription initiation sites.";
RL J. Biol. Chem. 266:3854-3862(1991).
RN [4]
RP SEQUENCE OF 34-55.
RC TISSUE=Lung;
RA MEDLINE=84051289; PubMed=6314908;
RT Iwata K., Blacher R., Soffer R.L., Lai C.Y.;
RT "Rabbit pulmonary angiotensin-converting enzyme: the NH2-terminal
RT fragment with enzymatic activity and its formation from the native
RT enzyme by NH4OH treatment.";
RL Arch. Biochem. Biophys. 227:188-201(1983).
RN [5]
RP FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF
RN THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE
RN VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.
CC -I- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide--I-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC Asp nor Glu. Converts angiotensin I to angiotensin II.
CC -I- COFACTOR: BINDS 2 ZINC IONS (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- ALTERNATIVE PRODUCTS: 2 isoforms; somatic (shown here) and testis-
CC specific (AC P22966); are produced by alternative splicing.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62551; CAA44428.1; -.
DR EMBL: M58579; AAA31151.1; ALT_SEQ.
DR PIR: A23455; A23455.
DR PIR: S35484; S35484.
DR MEROPS: M02.001; -.
DR InterPro: IPR001548; Peptidase_M2.
DR Pfam: PF01401; Peptidase_M2; 2.
DR PRINTS: PR00791; PEPTIDASEA.
DR ProDom: PD004184; Peptidase_M2; 2.

```

DR PROSITE: PS00142; ZINC\_PROTEASE; 2.  
 KM Hydrolyse: Metalloprotease: Carboxypeptidase: Zinc: Dipeptidase:  
 KM Glycoprotein; Transmembrane; Repeat; Signal: Alternative splicing.  
 FT SIGNAL 1 33  
 FT CHAIN 34 1310  
 FT DOMAIN 34 1263  
 FT TRANSMEM 1264 1280  
 FT DOMAIN 1281 1310  
 FT REPEAT 222 588  
 FT REPEAT 829 1185  
 FT METAL 395 395  
 FT ACT\_SITE 396 396  
 FT METAL 399 399  
 FT METAL 992 992  
 FT ACT\_SITE 993 993  
 FT METAL 996 996  
 FT CARBOHYD 59 59  
 FT CARBOHYD 79 79  
 FT CARBOHYD 151 151  
 FT CARBOHYD 323 323  
 FT CARBOHYD 449 449  
 FT CARBOHYD 513 513  
 FT CARBOHYD 681 681  
 FT CARBOHYD 699 699  
 FT CARBOHYD 718 718  
 FT CARBOHYD 946 946  
 FT CARBOHYD 1195 1195  
 FT CONFLICT 48 48  
 SQ SEQUENCE 1310 AA; 150405 MW; 04777FAB17981DEA CRC64;  
 Query Match 29.9%; Score 1283.5; DB 1; Length 1310;  
 Best Local Similarity 36.8%; Pred. No. 5.3e-82;  
 Matches 275; Conservative 135; Mismatches 262; Indels 75; Gaps 17;  
 7 LLSLVAVTAOSTI-----EEOAKTFLDKFNEHEDLFYOSLSANMYNTNI 54  
 21 LLLLLLPPPALTLDPGLPGDADEAGARLFASVSSAEQVLFSTASMAHDINI 80  
 55 TEENVOANNADKNSAFKEOSTLAQMTPLDET-----QNTL---VKILOALQOQSS 106  
 81 TAEANAROE-----EALISOFAFAMGKRAKELYDPVQNTDELRIITIGAVRTLP 135  
 107 VSEDESKRLNTIINTMSTYSTGKVCNPDNQECLELPGNEIMANSLOVNEKLMWE 166  
 136 NPLAKROQNSLSMSQIYSTGVCPFNKTAQSCWSDPDNLNLTASRSTAMLFAME 195  
 167 SWRSEYQQLRPLEYEYVLLKNEMARAHYEDYGDYRWGDEYVNGVDYDSRGQILIDV 226  
 196 GWHNAVGIPLKPLQOEFALSNFARODGFSPTGAYWRSWYDSPTFE-----BDL 245  
 227 EHTFEIRKLYEHLAYVAKIMANV-PSYISPIGCLPAHLIGDMWGRFWNLVSLTPF 285  
 246 EITIDLELYLNLHAYVAVHRRYGRXYINLRPIPAHLIGNMAQSWESITVYMPVF 305  
 286 GOKPNIDVADAVDOAMDARIFKEAEKFEVSVGPNNTOGFWEMSKMLDPGNOYKACH 345  
 306 PKRPLDVTSTWYOKGMNATHFRAVEEFTISGLPPRPEWASMLKEKPDGGEVYCH 365  
 346 PTAMD-LGKDFRILMCTKVTMDDELTAHENGHIQYDMAVAAOFPLLRNGANESGEHAV 404  
 366 ASAMDFYNNKDFRIGOCQVNTDQSLYHENGHYQYLYQYDQVSLRR-ANPEFHAI 424  
 405 GELMSISATPKYHLSIGLSPDOEDNTEINFLKQALITVGLPTPTYLEKRWYVF 464  
 425 GYVLAISTYPAHLKIGILD-HVYNDESDINYLKMALEKIAFLPGYUVDQWRGVF 483  
 465 KEIIPDOMKKMKMKREIVGVVPRHDEYCPASLEHNSNYSIRYTRTLXQFO 524  
 484 SGRITSSSRKNEDWYLRITYGQICPPVVRNETHFPAKAFHLPSTVPIRIFVSVLDFQ 543  
 525 FOEALCOAAKHGCPHLKCDINSTAGOKLFNMLRLGSEPTLLEVVGAKNNVAPPL 584

Db 544 FHOALCMGAEHQGPLOHCQDIQYSTRAGAKLRAVLQACSRPMQOELVKDMVASDALDAQPL 603  
 QY 585 LNYEPLEFTYWKDONKNS--FVGH-STDNSPYADQSIKAVISLSKALG-----DRA 632  
 Db 604 LDYFQPYQWLOEQENENGEVLGMPYQWRPPLNNPEGIDVLTPEAASRFEVEYDRS 663  
 QY 633 YE--NND-----NEMTLFRSSVAVARQYELKVKNMOMIFGEDVAVANLKPRISFNPF 684  
 Db 664 FOAVNNEAEANWYNNTNITTEASKILLQKNNQJANTHTLYG-----NWARFVSNF 716  
 QY 685 VTAP-----KNVSD-----IIPREVER 702  
 Db 717 QNATSKRIKKVQDLQRAVLPAKLEEE 743  
 RESULT 8  
 ACET\_RABBIT STANDARD; PRT; 737 AA.  
 ID ACET\_RABBIT STANDARD; PRT; 737 AA.  
 AC P22968;  
 DT 01-AUG-1991 (Rel. 19, last sequence update)  
 DT 01-AUG-1991 (Rel. 41, last annotation update)  
 DT 15-JUN-2002 (Rel. 41, last annotation update)  
 DE Angiotensin-converting enzyme, testis-specific isoform precursor  
 DE (EC 3.4.15.1) (ACE-T) (Dipeptidyl carboxypeptidase I) (Kininase II).  
 GN DCPI OR ACE.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=New Zealand white; TISSUE=Testis;  
 RX MEDLINE=9380303; PubMed=2550457;  
 RA Kumar R.S., Kusari J., Roy S.N., Soffer R.L., Sen G.C.;  
 RT "Structure of testicular angiotensin-converting enzyme. A segmental  
 RT mosaic isozyme."  
 RL J. Biol. Chem. 264:16754-16758(1989).  
 RN [2]  
 RP SEQUENCE OF 1-117 FROM N.A.  
 RX MEDLINE=91139683; PubMed=1847388;  
 RA Kumar R.S., Thekkumkara T.J., Sen G.C.;  
 RT "The mRNAs encoding the two angiotensin-converting isozymes are  
 RT transcribed from the same gene by a tissue-specific choice of  
 RT alternative transcription initiation sites."  
 RL J. Biol. Chem. 266:3854-3862(1991).  
 CC -1- FUNCTION: CONVERTS ANGIOTENSIN I TO ANGIOTENSIN II BY RELEASE OF  
 CC THE TERMINAL HIS-LEU, THIS RESULTS IN AN INCREASE OF THE  
 CC VASOCONSTRICTOR ACTIVITY OF ANGIOTENSIN.  
 CC -1- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,  
 CC oligopeptide-I-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither  
 CC Asp nor Glu. Converts angiotensin I to angiotensin II.  
 CC -1- COFACTOR: BINDS 1 ZINC ION.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; testis-specific (shown here) and  
 CC somatic (AC P12822); are produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: SPERMATOCYTES, ADULT TESTIS.  
 CC -1- INDUCTION: EXPRESSION IS THOUGHT TO BE SUBJECT TO HORMONAL  
 CC REGULATION BY ANDROGENS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J05041; AAA31153.1; -;  
 DR EMBL: M58580; AAA31152.1; -;  
 DR PIR: A34402; A34402.  
 DR MEROPS: M02.004; -;  
 DR Interpro: IPR001548; Peptidase\_M2.









```

Db 297 LKMFQWDDFFQSGMLKLEQEFWDKSLLEKPDGGRDLVCHASAMDFYLLDDVRIKQCTR 356
Oy 364 VTMDFTLTAHEHMGHIQYDANAYAAQPLLNGANESGFHEVNGIMSLSATPHTHISGL 423
Db 357 VTQOQFEVHEHMGHIQYDLOQHPVYRTGAPNGFHEVNGVLSLSTPHHLEAVGL 416
Oy 424 LSPDFQENEFNEINFLKQALTYGLPTMYLEKRWMEVFKGEIPRDMQMKWEMKRE 483
Db 417 LK-NVSDNEARINQLETLTADKIVFLPFAFTMDKYRMALFROADKSENCACFWLREB 475
Oy 484 IGVVEPVPHDETCDPASELFHVSNDYSFIRYTRTLYQFOFOALCOAA-----KHBC 537
Db 476 YGIEPPVVRTEKDFDAPAKYHVSADVEYLRIVSFILQFOFYKSACITGEVPPQTEX 535
Oy 538 PLHKCDISNSTEAGOKLFNMLRLKSEFWTTLLENVYGANMMVRLPLNFEPLTWLK 596
Db 536 PLDNCDDYGSKEAGKLFENMISLGSKFPDPALEAFNGERIMTKALAEFEPLRWLE 594

RESULT 11
YFWA_BACSU STANDARD: PRT: 501 AA.
AC P50848;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical metalloprotease yPWA (EC 3.4.24.-).
GN YFWA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignelli S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Enlian K.D., Errington J., Fadre C., Ferrari E., Galleron N.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi A.,
RA Guisepi G., Guy B.J., Haga K., Hatoch J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kunano M.,
RA Kuriita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,
RA Medina N., Mellado R.F., Mizuno M., Moesti D., Nakai S., Nodack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portollik S., Prescott A.M.,
RA Priescan E., Pujic C., Rocha E., Roche B., Rose M., Sadele Y.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadele Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Tosechi M., Tamakoshi A., Tanaka T., Terpsita P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
RA Viari A., Wambut R., Wambuto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zanchin A.;
RA "The complete genome sequence of the Gram-positive bacterium Bacillus

```

```

RT subtilis";
RL Mature 390:249-256(1997).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M32.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to license@isb-stb.ch).
CC
CC EMBL: L47838; AAB38482.1; -
CC EMBL: L77246; AAA96610.1; -
CC EMBL: Z99115; CAB14125.1; -
CC MEROPS: M32.0PW.
CC Subtilisin; B611458; yPWA.
CC Interpro: IPR001333; Peptidase_M32.
CC Interpro: IPR000130; Zn_M32peptidase.
CC Pfam: PF02074; Peptidase_M32; 1.
CC PRINTS: PR000998; CRBOXYPTASET.
CC PROSITE: PS00142; ZINC_PROTEASE; FALSE_NEG.
CC Hypothetical protein; Hydrolase; Metalloprotease; Zinc;
CC Complete proteome.
CC METAL 265 ZINC (CATALYTIC) (POTENTIAL).
CC METAL 266 POTENTIAL.
CC ACT_SITE 266 ZINC (CATALYTIC) (POTENTIAL).
CC METAL 269 ZINC (CATALYTIC) (POTENTIAL).
CC SEQUENCE 501 AA; 58174 MW; A7489ABEFA38F82 CRC64;

Query Match 3.1%; Score 135; DB 1; Length 501;
Best local Similarity 20.9%; Pred. NO. 0.053;
Matches 98; Conservative 65; Mismatches 164; Indels 142; Gaps 26;

Oy 180 YEEYVVL--KNEVA--RANHYEDYG-----DYRGDYVNGVDGYDS 218
Db 103 YKEVILCSAETFAWEAKGKSDFLSPYBLQLEFNKRETYV-----GYO-- 150
Oy 219 RGQIEDVEHTFEELKPLVYEHLLAYRAKLMNAPSYISPGICLPAHLIDGMGRFTNL 278
Db 151 -----EHPYALLDLFE--PGVTVYLD-----QLFREL 177
Oy 279 YSLVPE-----GKPNIDVTDAMVDQAMQRIEKAERFVSVGLPNTQGFWENS 331
Db 178 KEAIIPLVKQVYASGNP-----DVSFTKAPKKEKELSLYFQELGYDF--- 224
Oy 332 MLTPGNGVQAVCHPTAMDGLKGDPRILMCTKVTNDPLTA-----HHEMGHIQYD--MAX 385
Db 225 ---DGRIDETV--HPFATTLNRGDVAV--TTRVDEKDFRTALFGTTHGCHAIYEONIDE 278
Oy 386 AAQFPLLRNGANDEGFHA-----VGE-----IMLSAATPRHUKSIGLSPDF- 428
Db 279 ALSGTNISDGSMSGHSQSLEFNTIGRKNKHWTPIYKKIQRASVPYOFKDISL--DDEV 336
Oy 429 QEDNTEINFLIKQA--LTIVGLPTMYLEKRWMEVFKGEIPRDMQMKWEMKREIYG 486
Db 337 RAINSKRSFLRVEADELTLYPLHITIRYEIEK--ALFSNEVSIEDLPISLMNKYDYDG 393
Oy 487 VVEPVPHDETCDPASELFHVSNDYS--FIRYTRTLYQFOFOALCOQ-----AAKHEGPI 539
Db 394 I---TPQDADGIIQDVAHAGDGFYPSYALGTYVTAQLKQKMLDELPRFDALLERGER 450
Oy 540 HHCDDISNSTEAGOKLFNMLRLKSEFWTTLLENVYGANMMVRLPLNFEPLTWLK 597
Db 451 HPIK-----QWLEKVIHIGKRRKPLDIKDATG--EELNVAYLIDY 490

RESULT 12
GIGB_BACSU STANDARD: PRT: 627 AA.
AC P39118;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```
043 LANSVIAADQIFELVKKNOMLLFGEDDVRAVLKPRISENFYIAP 688
      | : | : | : | : | : | : | : | : | : | : | : | : |
531 F-SFIRYGOKH-----GMAVFYRQYD-----              |
```

DE 1,4-alpha-glucan branching enzyme (EC 2.4.1.18) (Glycogen branching  
GN enzyme).  
GN GLGB  
OC Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
RN NCBI\_TaxID=1423;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=168.  
RC MEDLINE=94195107; PubMed=8146641.  
RX Kiel J.A.K.W., Boels J.M., Beldman G., Venema G.;  
RT "Glycogen in *Bacillus subtilis*: molecular characterization of an  
RT operon encoding enzymes involved in glycogen biosynthesis and  
RT degradation";  
RL Mol. Microbiol. 11:203-218(1994).  
RL (2)  
RP SEQUENCE FROM N.A.  
RC STRAIN=168.  
RC MEDLINE=98048467; PubMed=9387221;  
RX Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;  
RT "Sequencing and functional annotation of the *Bacillus subtilis* genes  
RT in the 200 kb *rnb-dnaB* region.";  
RT Microbiology 143:3431-3441(1997).  
RN (3)  
RP SEQUENCE FROM N.A.  
RC MEDLINE=98044033; PubMed=9384377;  
RX Kunst F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,  
RX Azevedo V., Beterio M.G., Bessieres P., Bolotin A., Borchert S.,  
RX Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RX Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,  
RX Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RX Danicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RX Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RX Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RX Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandt G.,  
RX Giuseppe G., Guy B.J., Haga K., Haeche J., Harwood C.R., Henaut A.,  
RX Hillert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RX Joris B., Karmata D., Kashara Y., Kleier-Blanchard M., Klein C.,  
RX Kobayashi Y., Koelster P., Koningsstein G., Krogh S., Kumanu M.,  
RX Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RX Lee S.M., Levine A., Liu H., Masuda S., Maubel C., Medigue C.,  
RX Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RX Noone D., O'Reilly R.P., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RX Paresan E., Puig P., Portetelle D., Porwollik S., Prescott A.M.,  
RX Rieger M., Rivolta C., Rocha E., Rapoport G., Ray M., Reynolds S.,  
RX Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,  
RX Sekiguchi J., Sekowska A., Seror S.J., Serrero P., Shin B.S., Soldo B.,  
RX Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RX Takouchi M., Tanakoshi A., Tanaka T., Teplitsa P., Tognoni A.,  
RX Toesto V., Uchiyama S., Vandenbol M., Vannier F., Vasarotti A.,  
RX Viati A., Wadhut R., Wedler E., Wedler H., Weitznegger T.,  
RX Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,  
RX Yoshida K., Yoshikawa H.F., Zumshtein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
RT *subtilis*.";  
RL Nature 390:249-256(1997).  
CC -1- FUNCTION: CATALYZES THE FORMATION OF THE ALPHA-1,6-GLUCOSIDIC  
CC LINKAGES IN GLYCOGEN BY SCISSION OF A 1,4-ALPHA-LINKED  
CC OLIGOSACCHARIDE FROM GROWING ALPHA-1,4-GLUCAN CHAINS AND THE  
CC SUBSEQUENT TRANSFER OF THE OLIGOSACCHARIDE TO OTHER PARTS OF  
CC ALPHA-1,4-GLUCAN CHAINS.  
CC -1- CATALYTIC ACTIVITY: Formation of 1,6-glucosidic linkages of  
CC glycogen.  
CC -1- PATHWAY: Glycogen biosynthesis; third step.  
CC -1- INDUCTION: EXPRESSED EXCLUSIVELY ON MEDIA CONTAINING CARBON  
CC SOURCES THAT ALLOW EFFICIENT SPOULATION.  
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

RESULT 13
EP1B_STAEP STANDARD; PRT; 986 AA.
ID EP1B_STAEP
AC P30195;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Epidermin biosynthesis protein epib.
GN EP1B.
OS Staphylococcus epidermidis.
OC Plasmid PTU 32.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FU 3298 / DSM 3095;
RX MEDLINE=9215237; PubMed=1740156;
RA Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V.,
RA Goetz F., Entian K.-D.;
RT "Analysis of genes involved in the biosynthesis of lantibiotic
RT epidermin."
RL Eur. J. Biochem. 204:57-68(1992).
CC -1- FUNCTION: INVOLVED IN THE POSTTRANSLATIONAL MODIFICATION OF THE
CC LANTIBIOTIC EPIDERMIN.
CC -1- SUBCELLULAR LOCATION: POSSIBLY ASSOCIATED WITH, AND ANCHORED TO,
CC THE CYTOPLASMIC SIDE OF THE MEMBRANE.
CC -1- SIMILARITY: TO B.SUBTILIS SPAB AND L.LACTIS NISB.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X62386; CAA44253.1; -
RW Plasmid.
SQ SEQUENCE 986 AA; 116708 MW; 343CE98526847850 CRC64;
Query March 2.9%; Score 123; DB 1; Length 986;
Best Local Similarity 18.2%; Pred. No. 0.96;
Matches 153; Conservative 120; Mismatches 258; Indels 310; Gaps 40;
QY 1 MSSSSMLLSLVAVTAAGSTIEQAKTFLDKFNEH-----AEDL--FYOSSLAKMY 50
DB 320 ISEAVYIIMLLSPNHFGRKTRNVEHFFMDKYEFOALVNLKOLLSDINGFGYPRKDSYSF 379
QY 51 NNTNTEENVQNNNAGDKMSAFLEKQOSTLAOMYPLQETQNTLVKIQLOALQONSSVLSL 110
DB 380 SNNI-----AFLEK-----YLLAIQNNSHIETTE 404
QY 111 DKSKRLNTLNTMTSTYTGKVCNDPNQECLELPGLEIMANSLSIDNEELMAWESWRS 170
DB 405 NDVAKLEK-NNTVSKINA-----PVSTET-----YSELYF----- 433
QY 171 EYVKQLRPLYEYVVLKEMARANYEDYGDYWGCDYEVANGVDYSGQIIEDEVTET 230
DB 434 --GNSIKG-YEDFAVISPLGSGFNAGATGFR--TGNNFIK-----KKNOLOKEIVH-- 480
QY 231 EEIKRLYEHLHAYVRAKLMNAVPSYISPLGCLPAHLGLDMGRFNTNLSLTVPGGKPN 290
DB 481 -----HYNNYNNENGLETISQLNEGPLNSRNVLNN--NRIYNTCLNLP--KSD 527
QY 291 IDVTDAWDADQADRIKE--AAKFFVSGLPMMTQG-----FWNSMLTDPG 337
DB 528 IDINIFIGATFNKLYISEKSDSLVFSNSGMFVEGSELXKFLREISFEKTFIQPI 587
QY 338 N-----VQKAVCHPTAMDGLGKGFRLIMCTKTYMDPLFAHHEMGHIQYDM 383
DB 588 TEEGIDSLPFCPRIIYKNIILKPAWKINSEMFSE--TENMLNRFAT--REKWHIPADV 643

```

```

QY 384 AYA-----AOFELLRNGANGFHEAVEIMS----- 409
DB 644 IIAFGDNRLNLNDKHLILKKEKKHGRIRILESTINESNRMELIYTPLYKKTSL 703
QY 410 --LSAATK-----HLK---STGLSPDEQDENETENPLKALITVGLPTPTYLE 457
DB 704 KQOSFTIPKRNKRNKHNKDMFSIHLSIPKTYQD-----NFIDYLLPFTLEKVNPN 758
QY 458 KRWAVYKKEIKQDMKMKWEMKREIYGVPEVPHDEYCDPASLPHVSNDFSRFY- 516
DB 759 KEFYIKFKED--EDFTK--LRLLRE-----DEY-----SQIYSIRKWK 794
QY 517 -----TRITYOPFOFALCOAAKHGEPPLHKDISNSTAGOKLFNMLRGSEPTALE 571
DB 795 DYCLNSELXYDSIYDYVEYRYGGPHYEDIEENFPWDSLL--SINIQSE-FKIPKE 851
QY 572 NVYGAKNMVRPLNTFEPLFTWLDQKNKSVGMSDTMSPFADSIYRISKLALQDK 631
DB 852 FIYA--ISIDFLDYLE-----INSEKEEILINNA--ED 882
QY 632 AYEMNDNEMYLFRSSVAVAMROYFLKVNQMLLEGEDYRVANLKPRISFNFFVAPKNV 691
DB 883 LYRSND-----IREY-----KNLAKL-----INPKND 905
QY 692 SDIIP-----RTEVEKAIKRSRINDAF--RINDNSLEFLGIQ 728
DB 906 YELKKEPNLEHFLFNKISILENKLKTLQKSLYTSRSISGFIHMRGN--RIFGIN 961
QY 729 P 729
DB 962 P 962

```

```

RESULT 14
BGAL_BACME STANDARD; PRT; 1034 AA.
ID BGAL_BACME
AC 052847;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Beta-galactosidase (Ec 3.2.1.23) (Lactase).
GN BGAM.
OS Bacillus megaterium.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1404;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 319;
RA Strey J.;
RL Submitted (SEP-1997) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC galactose residues in beta-D-galactosides.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ000733; CA04267.1; -
DR HSSP; P00722; IBGL.
DR InterPro; IPR004200; Bgal_small_C.
DR InterPro; IPR004199; Bgal_small_N.
DR InterPro; IPR001649; GH_2.
DR Pfam; PF00703; Glyco_hydro_2_1.
DR Pfam; PF02836; Glyco_hydro_2_C_1.
DR Pfam; PF02837; Glyco_hydro_2_N_1.
DR Pfam; PF02929; Bgal_small_N_1.
DR Pfam; PF02930; Bgal_small_C_1.

```

DR PRINTS; PRO0132; GLHYDRLASE2.  
 DR PROSITE; PS00719; GLYCOSYL\_HYDROL\_F2\_1; 1.  
 DR PROSITE; PS00608; GLYCOSYL\_HYDROL\_F2\_2; 1.  
 KW ACT\_SITE 481 481 PROTON DONOR (BY SIMILARITY).  
 FT ACT\_SITE 547 547 NUCLEOPHILE (BY SIMILARITY).  
 SEQUENCE 1034 AA; 118673 MW; 38644C9A649415E9 CRC64;

Query Match 2.8%; Score 121; DB 1; Length 1034;  
 Best Local Similarity 19.2%; Pred. No. 1.4; Indels 206; Gaps 30;  
 Matches 104; Conservative 66; Mismatches 165; Indels 206; Gaps 30;

QY 29 LKFNHEDLFFYOSSL-----ASWNYNTITEENVQNMN---NAGDKSGLFKE 75  
 DB 260 LNFYFHDQDLTFEVLVLDANAQEVLDQPLQTNLSVSDQRTSLTFHFKSPAKMSA----- 315  
 QY 76 QSTLQMPYLOLQNL-IVKLDLQALQNGSSVLEDSK---RLNTLINTSTYTGK 131  
 DB 316 -----ESPMLYLYVL---SLKNAAGSIITETESCKVGFRTFEIKNGIMLT--NGK 359  
 QY 132 -----VCPNPPOECILLEPGLNEMANSIDYNERLM----- 163  
 DB 360 RVLQGVNHERDYSKRGAGITREDMIDILMKQHINAVTSHYPMDSVWYELCNEYG 419  
 QY 164 -----AN-----ESMRSEVGKQLNPLYE----- 182  
 DB 420 LVYIDETMLETHGTWTLQEGOKAVPGSKPEMKENVLDRCRSMYERDKNHPSTIIISLG 479  
 QY 183 -----YVVLK-NEMARANHEDYGDYMGDEYVNGCYDYSKROLIEDVHT 229  
 DB 480 NESFGENFQHMHTFFEKEDSTRLVNHE--GIFFHNRDASDIETMYVKP--ADVE-- 532  
 QY 230 FEELKPLVYENHLAVYRAKLNAPSYISPIGCLPAHLIGDMWGFRTVLYSLVPEGQK 289  
 DB 533 -----RYALMPKPKXYIL--CEXSHANGSCG---NLVYVWELFDQY 570  
 QY 290 NID---VTDAMVDQMDQRIFFKKE--KEFVSVG-----LPNTQGFWENSMKLTDPGNV 339  
 DB 571 ILQGGFIWD-----WKDQALQATAEDETSTYLAAGGPGDPN--DGNCCNGLIFFADGTA 623  
 QY 340 QKAV-----CH-PTAW---DLGKGFRLMCTKVTYMDFLTAHENGHI-----QYDMAY 385  
 DB 624 SPKIAEYKCKYQPVKWTAVDPAKGFV-----QNKHLFTNLNAYFPVW 667  
 QY 386 AAQPEPLRNGANEGFHEAVGEINSLSATPRHLKSGIGLSPDFQEDNETEINFLKQALT 445  
 DB 668 IVE-----KNG-----ELVEKHASLNLNAPDGTDELFLSYPLVYQENMETD-EFVLTLSTR 716  
 QY 446 I 446  
 DB 717 L 717

RESULT 15  
 OSH2\_YEAST 2.8%; Score 119.5; DB 1; Length 1283;  
 ID OSH2\_YEAST STANDARD; PRT; 1283 AA.  
 AC Q12451; P89891;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE Oxyterol-binding protein homolog 2.  
 GN OSH2 OR YDLO19C OR D2845.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c;  
 RA Ureastarazu L.A., Andre B., Visers S.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SUBCELLULAR LOCATION.

RX MEDLINE=21301806; PubMed=11408574;  
 RA Levine T.P., Munro S.;  
 RT "Dual targeting of Oshp, a yeast homologue of oxysterol-binding  
 RT protein, to both the Golgi and the nucleus-vacuole junction.";  
 RL Mol. Biol. Cell 12:1633-1644(2001).  
 RN [3]  
 RP GENETIC ANALYSIS.  
 RX MEDLINE=21135676; PubMed=11238399;  
 RA Ben C.T., Cool L., Phillips J., Rine J.;  
 RT "Overlapping functions of the yeast oxysterol-binding protein  
 RT homologues.";  
 RL Genetics 157:1117-1140(2001).  
 CC -1 SUBCELLULAR LOCATION: Cytoplasmic. When bound to oxysterols it  
 CC translocate to the periphery of Golgi membranes.  
 CC -1 SIMILARITY: BELONGS TO THE OSBP FAMILY.  
 CC -1 SIMILARITY: CONTAINS 2 ANK REPEATS.  
 CC -1 SIMILARITY: CONTAINS 1 PH DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).

QY 10 SLVAVTAOSTIEBQAKTFLDKFNHEDLFFYOSSLASW---NYNTITEENVQ---NMN 63  
 DB 412 ALISHSKTQGSLEPESAOYQHTLKE---VIGPSSVSYLRPSNNLSVSSSEIQLNQMLT 468  
 QY 64 NAGDKSGLFKEOSTLAQYV-----LQEIQLVTKLQALQNGSSVLE-DSK 114  
 DB 469 ESKRNVSAKMINRLDGSKTPGVHTGSAILOVRSSNT-LKSNRMOGSGVAPIDKVP 527  
 QY 115 RLNTLINTSTYST-----GKYCPNPPOECILLEPGLN-----EIMANS 155  
 DB 528 NCANLSQSWTTTGTSASLSDNNYINDFEGDEANSDEEDL---GTFPDEEYTKQY 583  
 QY 156 LDYNERLAWESWRS-----EVGKQLRPLYEEVVLKNEMAR----- 192  
 DB 584 GPYKKEKLDYEQALISIELSLIELQEGEPSEVWLTKKSLINTSTFGKLDYTKRD 643  
 QY 193 -----ANHYEDYGDYMGDEYVNGVDYDSKROLIEDVHTFEELKPLYEHLHAYVRA 246  
 DB 644 KRLVDVMSKQGVNVVW-----VQSVKLELMELSKTERLASIDKERRGLKKIHL---K 694  
 QY 247 KLNAPYSYISPIG-----LPAHLIGDMWGFRTVLYSLVPEGQKPNIDYTDAM 297

```

DB 695 KLEESATATGKNSLENDKEQSDTASTTGQI-----AKETISAT-----KEDESDA- 743
DB 298 VDOAMPDAORIFKEAEKEFEVSVGLPNTQGWENSMITDPCNVOKAYCHPTAMDGLGKDR 357
DB 744 -DEFYDAAEVLDE-----VTELETAHPELSTAAKHPAPPV--PNETD----- 784
DB 358 ILMCTKVTMDFLTAHHEMCHIDYDAAQAOPLLRNGANEHGVAGELMISATPKH 417
DB 785 -----NDQYVODEKSIENVEKTSQKEKONNL-----VTEDEPKTDS 825
DB 418 LKSTIGLSPPOEDNEEINFLKO-----ALTIYGLPTFLYMLEKRWANVK 465
DB 826 LKNEK-----ADEKESQVKEKTEKLASSVIGEKTYAVTTVKKREYILKKGSTIGYE 879
DB 466 GEIPKDOMMK-----WMEKREIYV-----VPEVPDDETYCDPASLPHVNDYSFL 513
DB 880 DGIKRLSMDDKDPKISLMAVILKSMYKGMKMTWTLTPV-----TFNEPTSLQ----- 927
DB 514 RYTRITLYQFQFQALCOAKHE-----GPIHKCDISNTEAGQKLEFNLRLG--- 561
DB 928 ---RVADELEYSELDOAAATFEDSTLRTLYVAATASSVASTTKRYAKFPNPL-IGETP 982
DB 562 -----KSPWTLALENVYAK-----NMN 580
DB 983 EYSPDKQYRFTFOVSHHPPISTATWTESPRMDWGESEFVDTKENGSEFVNHGLMHIK 1042
DB 581 VRPLLYFEPLFTWLKDONK--NSFVGSWSDSPYADOSIKRISLSKALGD-----KA 632
DB 1043 LRPNDEKEELIYWKKNFTVIGILG-----NPOVDNHEEVN-VNHTTGHDCKLYEKA 1096
DB 633 YENDNEMLYFRSSVAAAMR-QYFLKYKNQMLLEGEEDVAVANLKRIPISNFEVTAPKNV 691
DB 1097 RGM-----RSSGALEYIGEYVNNKKQKWLIGH-----WKEAIFAKKV 1136
DB 692 SDIIPREVEKAIKRSRIND-----AFRLNDSLEFLGICP---TLGPPNPV 739
DB 1137 KD--GDSLEKT-RTASAGNGPDDGKFLIMKANRPREPFULPFPATILNAP-QPHL 1192
DB 740 SIWL 743
DB 1193 LPWL 1196

RESULT 16
YAMB_SCHPO STANDARD; PRT; 3655 AA.
ID YAMB_SCHPO STANDARD; PRT; 3655 AA.
AC Q10064;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein CLF5.11c in chromosome I.
GN SPAC1F5.11C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Squares J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown S., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hiderg J., Hodgson G.,
RA Holtby S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Muddett D., Odeli C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

```

```

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Gymnopiez B.,
RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Filz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Talaue V., Mottier S.,
RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado U., Jimenez J., Sanchez M., del Rey F., Benito G.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Uesery D., Bartell B.G., Nurse P.;
RA "The genome sequence of Schizosaccharomyces pombe."
RA Nature 415:871-880(2002).
CC - SIMILARITY: STRONG, TO YEAST YHR099W.
CC - SIMILARITY: BELONGS TO THE P13/P14 KINASES FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: Z68136; CAA92239.1; -.
CC InterPro: IPR003151; FAT.
CC InterPro: IPR003152; FATC.
CC InterPro: IPR000403; P13_P14_kinase.
CC InterPro: IPR001440; TPR.
CC Pfam: PF00454; P13_P14_kinase; 1.
CC Pfam: PF02259; FAT; 1.
CC Pfam: PF02260; FATC; 1.
CC SMART: SM00146; P13KC; 1.
CC PROSITE: PS50290; P13_4_KINASE_3; 1.
CC DR Hypothetical protein; Transferrase; Kinase.
CC FT DOMAIN 3324 3655 P13K/P14K (BY SIMILARITY).
CC SEQUENCE 3655 AA; 420774 MW; 50475E33FC2124A CRC64;

Query Match 2.8%; Score 119.5; DB 1; Length 3655;
Best Local Similarity 17.0%; Pred No. 11;
Matches 93; Conservative 93; Mismatches 197; Indels 165; Gaps 23;

DB 4 SSWLLSLVAVTAQSTIEBOAKT---FLDKFHEADELFY---OSSLASMYNTNTE 56
DB 2633 SDW-----STEDPLEKATKSLSPFSLRHRTADALLYLNKQKKGSVTEFSRIID 2684
DB 57 ENYONNMNADKSAFLKE--OSTLAQMPLOEIQNLTKLQALQALQNGSSVLSKDKK 114
DB 2685 ECKQFSLR--RWQQLPKRYQSHVSLHHEDELVEL-----QAFGITYSLNDT 2731
DB 115 RLNTILNTMSTIYSTGKVCNDPNDPDCILPEGLNEMANSIDYNERLMA-WESWRSEVG 173
DB 2732 NIHHIDKL-----RDIKVVLQGWRELPVWD-DIDIMSDILAMQSVF 2775
DB 174 KQLRPLVEEYVVLKNEKARNHEDYGDYKRGDEVNGVGYYSKGLLEDEHFFEL 233
DB 2776 KSINKVFLPLVLSIAQOSTNKSNTNSVYLRYGHEI----- 2811
DB 234 KPLEVNHAVYRAKIMAYPSYISPICLPAHLIGDMWGRFNTNLSTVFPQOKPNIDV 293
DB 2812 -----AMINRF-AHARVNHLEPVCINL-----IKITVL-----PNEI 2846
DB 294 TDAMVDOAMPDAORIFKEAEKEFEVSVGLPNT-----QGWENSMITDPR 337
DB 2847 QEATFLKREQECHYESPSEMOGLEVINNTNLMYFRRROKAEFTLKGFGON-RIGEKD 2905
DB 338 NVQAAVCHPTAMDGLGKDFRIMCTKYTMDFLTAHHEMCHIDYDAAQAOPLLRNGAN 397
DB 2906 EANOAFATVAOIDIGSG-----KAWSEMG-LIHDELFOANPQEIHNACN 2948
DB 398 --EGFHEAVGEIMSLSATPKHLKSLGSLSPFOEDNETELINLLKQALTIVGLPTTYM 455

```

Db 2949 AVSCFLOA-SSLLSSNSKPLTLRYLWLLVDSDHSGVSEV-----YSSFKSE 2995  
 Qy 456 LKRMVAVKGEIPKDDMMKKMKRELVGVPEVDEYCDPASLFHVSNDY-SFTR 514  
 Db 2996 IFTWNIIPF---IPQ-----LALSRESIHARAILIOIAKYPOSTLH 3036  
 Qy 515 YVTRTLQY 522  
 Db 3037 FOLRTAYE 3044

RESULT 17  
 DTHC\_PARTE STANDARD; PRT; 4540 AA.  
 ID DTHC\_PARTE STANDARD; PRT; 4540 AA.  
 AC Q27171;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Dyein heavy chain, cytosolic (DYHC).  
 GN DHC-8.  
 OS Paramacium tetraurelia.  
 CC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Penicillida;  
 CC Paramacium.  
 OX NCBI\_TaxID=5688;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Stock 51;  
 RX MEDLINE=96157890; PubMed=8589455;  
 RA Kandi K.A., Forney J.D., Asai D.J.;  
 RT "The dynein genes of Paramacium tetraurelia: the structure and  
 expression of the ciliary beta and cytoplasmic heavy chains.";  
 RL Mol. Biol. Cell 6:1549-1562(1995).  
 CC -1- FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A  
 MOTOR FOR THE INTRACELLULAR RETROGRADE MOTILITY OF VESICLES AND  
 ORGANELLES ALONG MICROTUBULES.  
 CC -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF  
 CC INTERMEDIATE AND LIGHT CHAINS.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U20449: AAA75445.1 -  
 DR InterPro: IPR004273; Dynein\_heavy.  
 DR Pfam: PF03028; Dynein\_heavy; 1.  
 KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.  
 FT DOMAIN 440 482 COILED COIL (POTENTIAL).  
 FT DOMAIN 698 722 COILED COIL (POTENTIAL).  
 FT DOMAIN 794 827 COILED COIL (POTENTIAL).  
 FT DOMAIN 975 995 COILED COIL (POTENTIAL).  
 FT DOMAIN 1169 1251 COILED COIL (POTENTIAL).  
 FT DOMAIN 1295 1311 COILED COIL (POTENTIAL).  
 FT DOMAIN 1933 1961 MICROTUBULE-BINDING (POTENTIAL).  
 FT DOMAIN 3076 3182 COILED COIL (POTENTIAL).  
 FT DOMAIN 3289 3367 COILED COIL (POTENTIAL).  
 FT DOMAIN 3653 3688 COILED COIL (POTENTIAL).  
 FT DOMAIN 3820 3851 COILED COIL (POTENTIAL).  
 FT DOMAIN 4238 4259 COILED COIL (POTENTIAL).  
 FT DOMAIN 4313 4342 COILED COIL (POTENTIAL).  
 FT NP\_BIND 1835 1842 ATP (POTENTIAL).  
 FT NP\_BIND 2129 2136 ATP (POTENTIAL).  
 FT NP\_BIND 2496 2503 ATP (POTENTIAL).  
 FT NP\_BIND 2834 2841 ATP (POTENTIAL).  
 SQ SEQUENCE 4540 AA; 528626 MW; 0C6103148B8B95F9 CRC64;

Query Match 2.8%; Score 119.5; DB 1; Length 4540;

Best Local Similarity 18.3%; Pred. No. 15;  
 Matches 159; Conservative 131; Mismatches 315; Indels 265; Gaps 46;

Qy 12 VAVTAOSTTEEOAK-----TFDGFNBAEPLF-----YSSLAS 47  
 Db 521 IAISEVATITTTNRDLSGASAKEMFRLAFLN---KLFRRPKAKAIQEQSOL-- 574  
 Qy 48 WNTYNTTEENVQNMNAGKWSAFLKEOST--LAQWPLQE--IQNTVYLOLQALQ 101  
 Db 575 ---LKTVDHKDIQSLQNKFE--TYQKSNRSLASARDIFLTSFVYWSQLOIRQKYM 628  
 Qy 102 QNGSSVL---SEPKSRLLTILNTMSTIYSTGKVCNPNPQCLLEGLNFIANSUD 157  
 Db 629 QKVEQIIGPQWAE-----TDGKKCKE-----METERILLD 660  
 Qy 158 YNEELMAMESRSEVG-----KQLPLVEEYVVLNEMARAHYEDYRGDEYVNG 211  
 Db 661 SGPAL---EDWQKEIHHNKAQVSNKLEPVYTRRGLEIRVYKEKLSQLK---EVAN 714  
 Qy 212 VDG---YDYSQQLIEDVETFEETKPLYEHLAYVR-----AKIMNAPSYISP 258  
 Db 715 LSMNKTQVPYSISHIANDAKASYFALSQESLHTYIQITSQLNMAKSAKVALAKREVQ- 773  
 Qy 259 IGLPLAILGDWGRFMTNLVSTVPEGQKPNIDVTDA--VDQMDA--QRIK----- 309  
 Db 774 ---LQIQGPNYIMTKTQLO-PYVK---FTDKVELEQVANGINERIGQIESIC 822  
 Qy 310 EAEKFFVSVGLPMTQGFENSNLTDGPNQAVC-----HPYAMDGKGFRLMCT 362  
 Db 823 EAMKTCPPVSLD-----KLKDQEVISLGCNNNSNHIMQIDDK-QIESIILD 872  
 Qy 363 KVV---MDPL-----TAHHEKGIQYMAAOP-----FLIRGAN 397  
 Db 873 RYTVQKEMWQINQYOKIOERGVLNQTVAHEL-KLDDIIVDPPEVAKYFWQE--- 928  
 Qy 398 EGFHEAVGEIMSLSATPKHL-----KSIGLSPDQEDNETEIN---FLIKQALTVG 448  
 Db 929 -FHKMIGQICSLPRVYANRFNTIQNGTGPNGTORDLSTTIKINQOLKDAYSOIG 986  
 Qy 449 TLPFTYMLEKRWVFKGEIPDDMMKW---WEKRELVGVPEVPHDE----- 495  
 Db 987 QL---LE-----DMEQYQVQWLNQYSLMEID---IKQVEDILDDEKQWQM 1027  
 Qy 496 ---TYDDPSLFL---VSNDSIRIKYTTITLYQFOFQELCOAAHHEGPKLKD 543  
 Db 1028 LVDIQGRATFQNSTEEHFGAIIIDRYVQVYKINHKYDAMHKKEL----- 1073  
 Qy 544 ISNSTEAGKLFNMLRGKSEFWTLALENVYGA---NMNVRPLNTEFPLTWLKDQ 599  
 Db 1074 ---NHFGMKFGEQLRVN-----KNVTEKELKLKINFOLDLSIESTITIOED 1121  
 Qy 600 KNSFGWSDNSPYVD-QSIKVRISLK---SALGDKAYENDENMYLFRSSVAYAMROYF 655  
 Db 1122 K-KFGWASADIESFNKGQVLDRORYQYGDWLSFEQVEMQWOFKQIRSKKLSQSEEM 1180  
 Qy 656 LKVKQMLIFGEDEVYVNLKPRISFNFEVYAPKVVSLIR--TEVEKAIKMRSRIND 713  
 Db 1181 NNIGQKI---QDPRYINQOQLEIEQWKTSPKSGGCSPEAEQILKSLNQLISVOE 1236  
 Qy 714 AFLRLNDSLEFLGIQTPGPNQPVSTWL 743  
 Db 1237 KYEKSQAKKEILKMP---PTHQOKLVNLL 1263

RESULT 18  
 MTHR\_CAEEL  
 ID MTHR\_CAEEL STANDARD; PRT; 663 AA.  
 AC Q17693;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Probable methyltransferahydrofolate reductase (EC 1.5.1.20).  
 GN C06A8.1.



OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 CC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPRAIN-Bristol N2;  
 RA Leimbach D.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 CC CATALYTIC ACTIVITY: 5-methyltetrahydrofolate + NADPH.  
 CC methyltetrahydrofolate + NADPH.  
 CC -1 COFACTOR: FAD (BY SIMILARITY).  
 CC -1 PATHWAY: Folate metabolism.  
 CC -1 SIMILARITY: BELONGS TO THE METHYLENETERAHYDROFOLATE REDUCTASE  
 CC (EC 1.5.1.20/EC 1.7.99.5) FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: U39849; AAA81048.2; -  
 DR HSSP: P00394; 1A5T.  
 DR WormRep: C06A8.1; CE30593.  
 DR InterPro: IPR004621; Fadh2\_euk.  
 DR InterPro: IPR003171; MethylOf\_redctse.  
 DR Pfam: PF02219; MTHFR.1.  
 DR TIGRFAMs: TIGR00677; fadh2\_euk.1.  
 DR Oxidoreductase; Flavoprotein; FAD; NADP.  
 KW SEQUENCE 663 AA; 75486 MW; EA94806A2C3BC1CD CRC64;  
 SQ  
 Query Match 2.7%; Score 118; DB 1; Length 663;  
 Best Local Similarity 18.3%; Pred. No. 1.2; Indels 210; Gaps 30;  
 Matches 101; Conservative 77; Mismatches 163;  
 QY 305 ORFKAERF-----VSGLPN-----MTGG-----FWNSMLTPGKVNOK- 341  
 DB 63 ERLIDKQPFSLFFPRFVNGVPRFLERVERLSGSGSVFVDMVWH--MGSDPANVDRV 120  
 QY 342 -----AVCHPTAMDLGKD----- 355  
 DB 121 TSSSSIAASMLDYCGVDTMLHMTQYVKNADTLKHLKLEOKAKANGLSIALRGDLPPTL 180  
 QY 356 -----FRILMCTKVTMDLFT-----AHHEMGHIQYDAVAQAOPFLRNKANGFHEAV 404  
 DB 181 EDTHQFRALDMIRKREKGNFYSICAGYPLGHQAP--SYKADMLYKAKKADAGANFYI 239  
 QY 405 GEIMSLSAATPKHL--KSIGLSPDFQEDNETEINFLKQALITVGLPT--YMLEKRR 460  
 DB 240 TQLFPAERPEKFEVROCRELIGTQP-----ILGIMPIMGESIKRI 281  
 QY 461 WAFKGEIPKQWKKMKKEKREIVGVVEPRPHDETCDPASLFHNSNDYSFIRYTRTL 520  
 DB 282 AKLSOLEIP-----QHLLDLLEIKHD-----DDAVOKYTEKRC 315  
 QY 521 YQFOFOALCOAKHGEPLKHCDSNSTEAGQKLFNMLRLGKSEP-----WTLLENV 573  
 DB 316 IE-----MCRRL-----LDNQTASIHLYTMNRGSLREILKSLGKMKLBGDVY 359  
 QY 574 VGAKNMVVRPL--LNFEPFLTWLKDONKNSEVGSWSTDSPIYDQSLKVAISLSKALGDK 631  
 DB 360 FPMKNRSQHPIRCLSEVPRIT-W--SRPRSYITRTMDQPFNG----- 401  
 QY 632 AYENDNEMYLFRSSVAYA-MRQVELK-----VKN-----QMLTGF-----EDVRANLKP 678  
 DB 402 --RWGNS-----SSPAFGDVSSYILSNLTITVKNADRLAMFGANIESFEDVK-----R 447

QY 679 ISPNFVTAPKXVSD-----IIPREVEKAIMSSSRINDAF-RINDNSLEFLGIQPTL- 731  
 DB 448 VFNITITQAPN--ADGKVATVLEPTEALFQVPEFSLISEQLWYCNENGLITLVNQPSPVN 505  
 QY 732 GPNQPPVSIM 742  
 DB 506 GAPSTDPVIGW 516  
 RESULT 19  
 ID AKA9\_HUMAN STANDARD; PRT; 3911 AA.  
 AC 099996; 09U004; 09U03; 09Y6Y2; 014869; 043355; 094895; 09Y6B8;  
 DT 16-OCT-2001 (Rel. 40; Created)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 16-OCT-2001 (Rel. 40; Last annotation update)  
 DE A-kinase anchor protein 9 (protein kinase A anchoring protein 9)  
 DE (PRK9) (A-kinase anchor protein 450 kDa) (AKAP 450) (A-kinase anchor  
 DE protein 350 kDa) (AKAP 350) (hAKAP 350) (AKAP 120 like protein)  
 DE (Hyperin protein) (Yotiao protein) (Centrosome- and golgi-1-localized  
 DE PKA-associated protein) (CG-NAP).  
 GN AKAP9 OR AKAP450 OR AKAP350 OR KIA0803.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 4).  
 RC TISSUE=Brain;  
 RA MDLINE=98151389; PubMed=9482789;  
 RA Lin J.W., Wyszynski M., Madhavan R., Sealock R., Kim J.U., Sheng M.;  
 RT "Yotiao, a novel protein of neuromuscular junction and brain that  
 RT interacts with specific splice variants of NMDA receptor subunit  
 RT NR1".  
 RL J. Neurosci. 18:2017-2027(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND VARIANT GLN-1347 INS.  
 RC MEDLINE=99219864; PubMed=10202149;  
 RA Witczak O., Skaalheg B.S., Keryer G., Bornens M., Tasken K.,  
 RA Jansen T., Gerstlitz S.;  
 RT "Cloning and characterization of a cDNA encoding an A-kinase anchoring  
 RT protein located in the centrosome, AKAP450".  
 RL EMBO J. 18:1858-1868(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RC TISSUE=Brain;  
 RA MEDLINE=99287934; PubMed=10358086;  
 RA Takahashi M., Shibata H., Shimakawa M., Miyamoto M., Mukai H., Ono Y.;  
 RT "Characterization of a novel giant scaffolding protein, CG-NAP, that  
 RT anchors multiple signaling enzymes to centrosome and the golgi  
 RT apparatus".  
 RL J. Biol. Chem. 274:17267-17274(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Kemmer W.A., Deliss S., Schwarz U.;  
 RT "Cloning of Hyperion".  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 323-3911 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Gastric parietal cell;  
 RA MEDLINE=99115654; PubMed=9915845;  
 RA Schmidt P.H., Dransfield D.T., Claudio J.O., Hawley R.G.,  
 RA Trotter K.W., Milgram S.L., Golderring J.R.;  
 RT "AKAP350, a multiple spliced protein kinase A-anchoring protein  
 RT associated with centrosomes".  
 RL J. Biol. Chem. 274:3055-3066(1999).  
 RN [6]  
 RP SEQUENCE OF 1802-3876 FROM N.A. (ISOFORM 5).  
 RC TISSUE=Lymphoblast;  
 RA Hinds R., Sutterer C., Becker M., Hawkins M.;  
 RT Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [7]



FT	DOMAIN	1585	1659		COILED COIL (POTENTIAL).
FT	DOMAIN	1857	2455		COILED COIL (POTENTIAL).
FT	DOMAIN	2544	2561		COILED COIL (POTENTIAL).
FT	DOMAIN	2603	2776		COILED COIL (POTENTIAL).
FT	DOMAIN	3065	3092		COILED COIL (POTENTIAL).
FT	DOMAIN	3124	3470		COILED COIL (POTENTIAL).
FT	DOMAIN	3587	3689		COILED COIL (POTENTIAL).
FT	DOMAIN	3726	3730		COILED COIL (POTENTIAL).
FT	DOMAIN	203	292		POLY-LEU.
FT	DOMAIN	321	1010		GLN-RICH.
FT	DOMAIN	1846	1010		GLU-RICH.
FT	VARSPLIC	17	28		MISSED (IN ISOFORM 2 AND ISOFORM 3).
FT	VARSPLIC	1637	1642		OLOREI -> LARRD (IN ISOFORM 4).
FT	VARSPLIC	1643	3911		MISSED (IN ISOFORM 4).
FT	VARSPLIC	2175	2182		MISSED (IN ISOFORM 3).
FT	VARSPLIC	2175	2183		SADFOAKE -> Q (IN ISOFORM 6).
FT	VARSPLIC	2895	2907		VEGFHMGFCFLC -> GSSPELAHSAPVOTREISS
FT	VARSPLIC				(IN ISOFORM 2, ISOFORM 3 AND ISOFORM 6).
FT	VARSPLIC	2895	2948		MISSED (IN ISOFORM 5)
FT	VARSPLIC	3901	3911		STOFHGMR -> ALSLTTSQWHSARPTAPFETLISH
FT	VARIANT	1347	1347		SLG (IN ISOFORM 6).
FT					K -> KQ.
FT	CONFLICT	76	76	/FtId=VAR_010926.	E -> Q (IN REF. 3).
FT	CONFLICT	475	475		M -> I (IN REF. 3).
FT	CONFLICT	554	554		E -> G (IN REF. 3).
FT	CONFLICT	638	638		R -> S (IN REF. 3).
FT	CONFLICT	663	663		N -> S (IN REF. 3).
FT	CONFLICT	913	913		H -> N (IN REF. 3).
FT	CONFLICT	956	956		K -> N (IN REF. 3).
FT	CONFLICT	980	982		OKH -> PKR (IN REF. 1 AND 2).
FT	CONFLICT	997	997		Q -> P (IN REF. 1 AND 2).
FT	CONFLICT	1001	1001		Q -> P (IN REF. 1 AND 2).
FT	CONFLICT	1020	1020		N -> D (IN REF. 3).
FT	CONFLICT	1028	1028		V -> E (IN REF. 3).
FT	CONFLICT	1626	1626		R -> P (IN REF. 1 AND 2).
FT	CONFLICT	1703	1703		N -> T (IN REF. 3).
FT	CONFLICT	1707	1707		V -> G (IN REF. 3).
FT	CONFLICT	1802	1803		MISSING (IN REF. 5).
FT	CONFLICT	1843	1843		A -> P (IN REF. 3).
FT	CONFLICT	1956	1956		I -> V (IN REF. 3).
FT	CONFLICT	2027	2027		V -> D (IN REF. 5).
FT	CONFLICT	2157	2158		EI -> HE (IN REF. 7).
FT	CONFLICT	2169	2169		E -> V (IN REF. 3).
FT	CONFLICT	2514	2514		L -> R (IN REF. 3).
FT	CONFLICT	2851	2851		I -> N (IN REF. 8).
FT	CONFLICT	2957	2957		E -> D (IN REF. 3).
FT	CONFLICT	2983	2983		P -> S (IN REF. 3).
FT	CONFLICT	3087	3087		Q -> H (IN REF. 3).
FT	CONFLICT	3218	3218		Q -> H (IN REF. 3).
FT	CONFLICT	3307	3309		ESR -> OSO (IN REF. 3).
FT	CONFLICT	3751	3751		P -> A (IN REF. 3).
FT	CONFLICT	3833	3833		T -> S (IN REF. 3).
SEQ	SEQUENCE	3911 AA;	453664 MM;	3FB1CB1C819B47AA CRC64;	
<hr/>					
Query Match					
Best Local Similarity		2.7%;		Score 118; DB 1; Length 3911;	
Matches		70; Conservative		63; Mismatches 130; Indels 84; Gaps 15	
OY	13 AVTAAGSTIEQKTFIDKFENHEAEELFYROSS-----LASMYNNITTEENVQNANNAAGDK 68	:::::	:~::~:	:	11
Dd	1095 SLRTGPSENDKIQKEINLVKSPQNDLRIDMAEQRICISLVSITHVOVR-EYMEKKRK 1153	:::::	:~::~:	:	11
OY	69 WSAFLKQSTLAOMYPLOETONL-TVUKIQQLAQONG-----SVLSDEKSK 114		:::::	:~::~:	11
Dd	1154 ALCSIKELLFAOEKRIKELQHDELQIMTKTGTOBGEGKRPLHLIGLKQAKAVSECSY 1213		:::::	:~::~:	11
OY	115 RUNTIMTMTIYSIRKVC-----RPDPQECLL--LEPGINTMANSLDYNE----- 160	:::	:~::~:	:	11
Dd	1214 PLQTGVSVGEYYTPALCKCEVAENAEKNSCYISENEDPELOYDRYEVQDFQENMHITLN 1273	:::	:~::~:	:	11
OY	161 -----RLAWAES--WNSVEGKQLRPIYE-EYIVVLKNEMARANHIEDYG 200	:::	:~::~:	:	11

DB 1274 KYTEENKLVLTQRLSKIMGQTDGKMLFEGEENLPKETEFLSIHSGMNTLE----- 1327

QY 201 DYWRDYEYNGVDYDYSRGOLIEVEHT-FBEIKPLYEHLHAYRAKLANAYSYSPI 259

DB 1338 -----DIVN-----HKSSLSIDPLEKTEEOVELESLSISQOOLKETEONYEAEI 1377

QY 260 GCLPAHLGDMWGRFMTNLYSLTFEGQKPNIDVTDAVDAMPOR 306

DB 1378 HCLQRLQA-----VSESTVP-----PSLPDVSIVTES-DAOR 1410

RESULT 20

UTRO\_HUMAN STANDARD; PRT; 3433 AA.

AC P46939:

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Utrrophin (Dystrophin-related protein 1) (DRP1) (DRP).

GN UTRN OR DMDL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=93096045; PubMed=1461283;

RA Tinsley J.M., Blake D.J., Roche A., Fairbrother U., Riss J.,

RA Byth B.C., Knight A.E., Kendrick-Jones J., Suthers G.K., Love D.R.,

RA Edwards Y.H., Davies K.E.;

RT "Primary structure of dystrophin-related protein.";

RL Nature 360:591-593(1992).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 144-261.

RA MEDLINE=99141377; PubMed=9887274;

RA Keep N.H., Norwood F.L.M., Moores C.A., Winder S.J.,

RA Kendrick-Jones J.;

RT "The 2.0-A structure of the second calponin homology domain from the

RT actin-binding region of the dystrophin homologue utrophin.";

RL J. Mol. Biol. 285:1257-1264(1999).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 31-256.

RA MEDLINE=20113481; PubMed=10647184;

RA Keep N.H., Winder S.J., Moores C.A., Walke S., Norwood F.L.M.,

RA Kendrick-Jones J.;

RT "Crystal structure of the actin-binding region of utrophin reveals a

RT head-to-tail dimer.";

RL Structure 7:1539-1546(1999).

CC -1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE

CC PLASMA MEMBRANE (BY SIMILARITY TO DYSTROPHIN).

CC -1- SUBCELLULAR LOCATION: NEUROMUSCULAR JUNCTION.

CC -1- TISSUE SPECIFICITY: MUSCLE.

CC -1- SIMILARITY: STRONG, TO DYSTROPHIN.

CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY

CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,

CC ABP-120, ABP-180, OR BETA-FODRIN).

CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.

CC -1- SIMILARITY: CONTAINS 20 SPECTRIN REPEATS.

CC -1- SIMILARITY: CONTAINS 1 ZK-TYPE ZINC FINGER.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X69086; CAA4829.1; -

DR PIR: S28381; S28381

DR PDB: 1BHD; 16-FEB-99.

DR PDB: 1QAG; 01-JAN-00.

DR Genew; HSCN:12635; UTRN.

DR MIM; 128240; -

DR InterPro; IPR001589; Actbind\_actnln.

DR InterPro; IPR001715; Calponin-like.

DR InterPro; IPR002017; Spectrin.

DR InterPro; IPR003122; Tsrh.

DR InterPro; IPR002349; Ww.

DR InterPro; IPR001202; Ww\_Rsp5\_WMP.

DR InterPro; IPR000433; Znf\_ZZ.

DR Pfam; PF00397; CH; 2.

DR Pfam; PF00435; spectrin; 19.

DR Pfam; PF00569; ZF\_1.

DR PRINTS; PR00403; WMDOMAIN.

DR SMART; SM00033; CH; 2.

DR SMART; SM00150; SPEC; 18.

DR SMART; SM00319; Tsrh; 1.

DR SMART; SM00456; Ww; 1.

DR SMART; SM00291; ZNF\_ZZ; 1.

DR PROSITE; PS00019; ACTININ\_1; 1.

DR PROSITE; PS00020; ACTININ\_2; 1.

DR PROSITE; PS50021; CH; 2.

DR PROSITE; PS01159; WW\_DOMAIN\_1; 1.

DR PROSITE; PS50020; WW\_DOMAIN\_2; 1.

DR PROSITE; PS01357; ZF\_ZZ\_1; 1.

DR PROSITE; PS0135; ZF\_ZZ\_2; 1.

KW Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;

KW Repeat; 3D-structure; Zinc-finger.

FT DOMAIN 1 246

FT DOMAIN 31 135

FT DOMAIN 150 252

FT REPEAT 253 308

FT REPEAT 309 417

FT REPEAT 418 526

FT REPEAT 541 637

FT REPEAT 687 798

FT REPEAT 803 902

FT REPEAT 1016 1083

FT REPEAT 1125 1230

FT REPEAT 1248 1334

FT REPEAT 1432 1541

FT REPEAT 1544 1649

FT REPEAT 1652 1753

FT REPEAT 1910 1968

FT REPEAT 1976 2081

FT REPEAT 2258 2333

FT REPEAT 2399 2440

FT REPEAT 2443 2556

FT REPEAT 2559 2636

FT REPEAT 2658 2688

FT REPEAT 2691 2797

FT DOMAIN 2812 2845

FT ZN\_FING 3064 3111

SO SEQUENCE 3433 AA; 394488 MW; EAE8DB409F858E5B CRC64;

Query Match 2.7%; Score 115.5; DB 1; Length 3433;

Best Local Similarity 18.5%; Pred. No. 19;

Matches 139; Conservative 98; Mismatches 249; Indels 265; Gaps 36;

QY 29 LDFEHEAEEL-----FYOSSLASWNYMTNTEENVQMNNAAGDKWSAFLEQSTL 79

DB 2533 LDDMNRWMDLAKKSSSTAHLEBASAKWNRLMSLEELIKWLNKDE---LKKOMPI 2588

QY 80 AOMYPLQEIQNTLVKQLQALQOQSSVLSDEKSKRLNTILMTSTIYSTGKVCNPDNQ 139

DB 2589 GGDVPAIQGYDCKALRELKEKREYSVLNAVDAQRY-----FLADQPIEAPEER 2639

QY 140 ECLLEPGL-----NEIMANSLDYNEELAMES-----WRSEVGQQLRPLYEE 182

DB 2640 RNLQSTTELTPEERACKIAKARKOSSEYKER--WESLNAVTSWMOQVAKALEKLDL 2696

QY 183 YVVLKEMARAHYEDYGDYMR--GDYEYNGVDGYDSGOLIEVEHTFEETIKPLYEHL 240

Db 2697 OGAMDDLDADMKAEBSVRNGMKPVGDLLIDSLQDH-----IEKIMAFREHIAPI----- 2745  
QY 241 HAYVRAKIMNAVPSYISPGCLP-----AHLLGDMGGRWTNLSLTPFGOKPNIDVTD 296  
Db 2746 --NFVKYKTVNDSSQSLSPDLHPSLKMSQDLDDLNR--W-KLLQVSV-----DD 2790  
QY 297 WYDQAMDAQRIKFAEKFFVSIVL-----PNNTOGF 328  
Db 2791 RLKQLQEAHRDGFSSQHLSTSVQLPWRORISHNKVPYIYNHOTQTCMDHPKKTLEL- 2849  
QY 329 ENSMLTDPGNV-----OKAVC-----HPTAMDLGKGF--- 356  
Db 2850 --QSLADLNNVRFSAVPTAIKIRRLQKALCLDLELSTNEIFKQH---KLNQNDQLLS 2903  
QY 357 --RIIMCTVYTMDDFLTAHMEGHQ-----YDMATYAQPFLLRNGANEGFH 401  
Db 2904 VPDVINCILTTTYDGLQMHKDLVNWPLCVMCLNMLNLYDTG-----RTGKIRVOS 2955  
QY 402 EAVGEIMSL-----AATPKHL---KSIGLLSPD----- 427  
Db 2956 LKIG--LMSLSKGLLEEKYRLEKVAAGPTKMDQROLGILLHDAIQIRQLGEVAFGGS 3014  
QY 428 -----FOEDNETEINFLKQALTYGLPTYMLEKMRMVFKEBIPROQMKRW 478  
Db 3015 NIEPSVRSCEFQNNN-----KPEISYKEFI-DWM 3042  
QY 479 EKKREIYGVVEPVPHDETYCDPA---SLEHVSNDYSFIRYTRTLYOFOEALCOAKH 535  
Db 3043 HLEPQSM-VMLPVLRVAAAEATAKHQAKCNICECPYIGFRYSILKHFNID--VCQS--- 3096  
QY 536 EGPRLHKCDISNSTEAGQKL-FNMLRL-----GKS-EPWTLALENVYGAKNMNR-PILN 586  
Db 3097 -----CFPSGRFAGHKLHYPMVEICPTSGEDVRDFTKVLKNKFRSKKYFAKHPRIG 3150  
QY 587 YEEPLFTWLKDQKNSFVGWSTDMSPYADQS 617  
Db 3151 YL-PVQTVLEGDNLFTPIPLISMWPEHYDPS 3180

Search completed: March 13, 2003, 16:57:36  
Job time : 43 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 16:53:52 ; Search time 36 Seconds

(without alignments)  
4607.442 Million cell updates/sec

Title: US-09-978-385-2

Perfect score: 4291

Sequence: 1 MSSSSWLLSLIVAVTAQAOST.....ISKGNPNQFQNTDVTQTSF 805

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPRENBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4291	100.0	805	4 Q9NRA7	Q9NRA7 homo sapien
2	4288	99.9	805	4 Q9NRA7	Q9NRA7 homo sapien
3	4281	99.8	804	4 Q9NRA7	Q9NRA7 homo sapien
4	3579	83.4	805	11 Q9NRA7	Q9NRA7 homo sapien
5	3485	81.2	798	11 Q9NRA7	Q9NRA7 homo sapien
6	1539	35.9	353	11 Q9NRA7	Q9NRA7 mus musculu
7	1336	31.1	732	6 Q9NRA7	Q9NRA7 mus musculu
8	1329	31.0	1304	6 Q9NRA7	Q9NRA7 pan troglod
9	1310	30.5	1313	11 Q9NRA7	Q9NRA7 pan troglod
10	1303	25.4	694	4 Q15540	Q15540 homo sapien
11	1089.5	25.4	648	5 Q9NRA7	Q9NRA7 mus musculu
12	1086	25.3	615	5 Q9NRA7	Q9NRA7 mus musculu
13	1057	24.6	660	5 Q17248	Q17248 boophilus m
14	1030	24.0	630	5 Q9VLAJ6	Q9VLAJ6 drosophila
15	1028	23.4	249	11 Q9D836	Q9D836 mus musculu
16	1004	23.4	249	11 Q9D836	Q9D836 mus musculu

17	930	21.7	844	5 Q8SXX2	Q8SXX2 drosophila
18	841	19.6	792	5 Q9VJY1	Q9VJY1 drosophila
19	642.5	15.0	907	5 Q18581	Q18581 caenorhabditis
20	502.5	11.7	661	5 Q9V520	Q9V520 drosophila
21	476	11.1	611	5 Q9VJY2	Q9VJY2 drosophila
22	434.5	10.1	202	11 Q64603	Q64603 ratius norv
23	409.5	9.5	628	5 Q9W021	Q9W021 drosophila
24	408.5	9.5	628	5 Q9W021	Q9W021 drosophila
25	387	9.1	222	11 Q9ESG3	Q9ESG3 ratius norv
26	380	9.0	222	4 Q9HJ8	Q9HJ8 homo sapien
27	386	9.0	222	11 Q9ESG4	Q9ESG4 mus musculu
28	306.5	7.1	121	6 Q9BDG1	Q9BDG1 bos taurus
29	280.5	6.5	135	4 Q16425	Q16425 homo sapien
30	251.5	5.9	157	11 Q61265	Q61265 mus musculu
31	239	5.6	775	5 Q9TX66	Q9TX66 drosophila
32	157	3.7	532	16 Q9KRV0	Q9KRV0 bacillus ha
33	154	3.6	502	16 Q8Y616	Q8Y616 listeria mo
34	152.5	3.6	734	5 Q9NKE3	Q9NKE3 drosophila
35	147	3.4	54	6 Q9Y161	Q9Y161 canis fami
36	147	3.4	502	16 Q92AC3	Q92AC3 listeria in
37	139.5	3.3	987	16 Q8YWG6	Q8YWG6 anabaena sp
38	139	3.2	608	16 Q9PR80	Q9PR80 ureaplasma
39	136	3.2	611	16 Q9PPW8	Q9PPW8 ureaplasma
40	130.5	3.0	461	2 Q9K152	Q9K152 enterococu
41	127.5	3.0	3571	10 Q9SL27	Q9SL27 arabidopsis
42	127.5	3.0	3574	10 Q9AUB4	Q9AUB4 arabidopsis
43	127	3.0	779	5 Q9XYS8	Q9XYS8 dictyostell
44	125.5	2.9	600	16 Q8R670	Q8R670 fusobacteri
45	124.5	2.9	987	11 Q61636	Q61636 mus musculu

## ALIGNMENTS

### RESULT 1

ID	Q9NRA7	PRELIMINARY	PRT	805 AA.
AC	Q9NRA7			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Angiotensin converting enzyme-like protein (ACE-related			
DE	carboxypeptidase ACE2).			
OC	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LYMPHOMA;			
RA	Tipidis S.R., Hooper N.M., Hyde R.J., Christie G., Karan E.,			
RA	Turner A.J.;			
RT	"A Human Homolog of Angiotensin Converting Enzyme - Cloning and			
RT	Functional Expression As A Captopril-Insensitive Carboxypeptidase.,"			
RL	Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Donoghue M., Hsieh F., Baronas E., Godbout K., Gosselin M.,			
RA	Stagliano N., Donovan M., Woolf B., Robison K., Jeyaseelan R.,			
RA	Breitbart R.E., Acton S.;			
RT	"A novel ACE-related carboxypeptidase (ACE2) converts angiotensin I to			
RT	angiotensin-9.," (2000).			
RL	Cite. Res. 0:0-0(2000).			
DR	EMBL; AF241254; AAF78220.1; -			
DR	EMBL; AF291820; AAF9721.1; -			
DR	MEROPS; M02.006; -			
DR	InterPro; IPR001548; Peptidase_M2.			
DR	InterPro; IPR001330; Zn_MTpeptidse.			
DR	Pfam; PF01401; Peptidase_M2; 1.			
DR	PRINTS; PR00791; PEPTIDASEA.			
DR	PRODOM; PD004184; Peptidase_M2; 1.			
DR	PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.			
KW	Carboxypeptidase.			

SEQ SEQUENCE 805 AA; 92462 MW; 8EE6E0A93150E8 CRC64;

Query Match 100.0%; Score 4291; DB 4; Length 805;  
Best Local Similarity 100.0%; Pred. No. 3.5e-305;  
Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSSSSWLLSLVAVTAOSTIEQAKTELDKFNHEADLFYQSSLASMNNTNITEENVQ 60
   1 MSSSSWLLSLVAVTAOSTIEQAKTELDKFNHEADLFYQSSLASMNNTNITEENVQ 60
Db 1 MSSSSWLLSLVAVTAOSTIEQAKTELDKFNHEADLFYQSSLASMNNTNITEENVQ 60
QY 61 MNNAAGDKWSAFLEKOSTLAQMYPLQEIQNLTVKLOLQALQNGSSVLSDESKRLNTIL 120
   61 MNNAAGDKWSAFLEKOSTLAQMYPLQEIQNLTVKLOLQALQNGSSVLSDESKRLNTIL 120
Db 61 MNNAAGDKWSAFLEKOSTLAQMYPLQEIQNLTVKLOLQALQNGSSVLSDESKRLNTIL 120
QY 121 NTMSTIYTGKVCNPDNPOECLELLEPGNEIMANSLDNERLMAWSEKSEVKGOLRPLY 180
   121 NTMSTIYTGKVCNPDNPOECLELLEPGNEIMANSLDNERLMAWSEKSEVKGOLRPLY 180
Db 121 NTMSTIYTGKVCNPDNPOECLELLEPGNEIMANSLDNERLMAWSEKSEVKGOLRPLY 180
QY 181 EEEVVLKNEMARANHYEDYGDYWRGDYEVNGVDYDYSRGOLIEDVHTFEETKPLYEHL 240
   181 EEEVVLKNEMARANHYEDYGDYWRGDYEVNGVDYDYSRGOLIEDVHTFEETKPLYEHL 240
Db 181 EEEVVLKNEMARANHYEDYGDYWRGDYEVNGVDYDYSRGOLIEDVHTFEETKPLYEHL 240
QY 241 HAYYRAKLMANAYPSYISPIGCLPAHLIDMGGRFNTLYSLTVPGQKPNIDYTDAMVQ 300
   241 HAYYRAKLMANAYPSYISPIGCLPAHLIDMGGRFNTLYSLTVPGQKPNIDYTDAMVQ 300
Db 241 HAYYRAKLMANAYPSYISPIGCLPAHLIDMGGRFNTLYSLTVPGQKPNIDYTDAMVQ 300
QY 301 AMDAORIFKEAEKFFVSVGLPNNTOGFENSMULTDPGNVOKAVCHPTAMDLGKDFRILM 360
   301 AMDAORIFKEAEKFFVSVGLPNNTOGFENSMULTDPGNVOKAVCHPTAMDLGKDFRILM 360
Db 301 AMDAORIFKEAEKFFVSVGLPNNTOGFENSMULTDPGNVOKAVCHPTAMDLGKDFRILM 360
QY 361 CTKVMTDDELTAHHEMGHIQYDMAVAAOPFLLRNGANGFHEAVGEINSLSAATPKHLKS 420
   361 CTKVMTDDELTAHHEMGHIQYDMAVAAOPFLLRNGANGFHEAVGEINSLSAATPKHLKS 420
Db 361 CTKVMTDDELTAHHEMGHIQYDMAVAAOPFLLRNGANGFHEAVGEINSLSAATPKHLKS 420
QY 421 IGLLSPDPOEDNETEINFLKQALTYGTLPTMYLEKRWNVFEGELIPKQOMKKWEM 480
   421 IGLLSPDPOEDNETEINFLKQALTYGTLPTMYLEKRWNVFEGELIPKQOMKKWEM 480
Db 421 IGLLSPDPOEDNETEINFLKQALTYGTLPTMYLEKRWNVFEGELIPKQOMKKWEM 480
QY 481 KREIVGVEPVPHDETYCDPASLFHVSNDYSFTRYTRTLVYOFQOALCOAAKHEGPLH 540
   481 KREIVGVEPVPHDETYCDPASLFHVSNDYSFTRYTRTLVYOFQOALCOAAKHEGPLH 540
Db 481 KREIVGVEPVPHDETYCDPASLFHVSNDYSFTRYTRTLVYOFQOALCOAAKHEGPLH 540
QY 541 KCDISNSTEAGQKLFNNMLRLGKSEPTLALENVGAKNNVRLNTEFEPLFTWLKDQNK 600
   541 KCDISNSTEAGQKLFNNMLRLGKSEPTLALENVGAKNNVRLNTEFEPLFTWLKDQNK 600
Db 541 KCDISNSTEAGQKLFNNMLRLGKSEPTLALENVGAKNNVRLNTEFEPLFTWLKDQNK 600
QY 601 NSFVGSTDMSPYADOSIKVRISLSKSGADKAYEMNDNEMYLERSSVAAKROYFLKAVN 660
   601 NSFVGSTDMSPYADOSIKVRISLSKSGADKAYEMNDNEMYLERSSVAAKROYFLKAVN 660
Db 601 NSFVGSTDMSPYADOSIKVRISLSKSGADKAYEMNDNEMYLERSSVAAKROYFLKAVN 660
QY 661 QMILFGEEDVRANLKPRISFNEFVTAAPKNVSDIIPTEVEKAIKMSRSRIINDAFLRNDN 720
   661 QMILFGEEDVRANLKPRISFNEFVTAAPKNVSDIIPTEVEKAIKMSRSRIINDAFLRNDN 720
Db 661 QMILFGEEDVRANLKPRISFNEFVTAAPKNVSDIIPTEVEKAIKMSRSRIINDAFLRNDN 720
QY 721 SLEFLGIQPTLGPNOQPVSIWLVGVVGVYIVGVILFTGIRDRKKKKKASGENP 780
   721 SLEFLGIQPTLGPNOQPVSIWLVGVVGVYIVGVILFTGIRDRKKKKKASGENP 780
Db 721 SLEFLGIQPTLGPNOQPVSIWLVGVVGVYIVGVILFTGIRDRKKKKKASGENP 780
QY 781 YASIDISKGENNPGFQNTDDVOTSF 805
   781 YASIDISKGENNPGFQNTDDVOTSF 805
Db 781 YASIDISKGENNPGFQNTDDVOTSF 805

```

RESULT 2

Q9BYF1 PRELIMINARY; PRG; 805 AA.  
Q9BYF1;  
AC 01-JUN-2001 (Tremblrel. 17, created)  
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, last annotation update)  
DE ACB2.  
GN ACB2.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Komatsu T., Suzuki Y., Sugano S.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RP SEQUENCE FROM N.A.  
RA Suzuki Y., Watanabe M., Sugano S.;  
RT "Cloning, expression analysis and chromosomal localization of a novel  
RT ACE like enzyme."  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB046569; BAB40370.1; -.  
DR MEROPS; M02.006; -.  
DR InterPro; IPR001548; Peptidase\_M2.  
DR InterPro; IPR000130; Zn\_MTPeptidase.  
DR Pfam; PF01401; Peptidase\_M2; 1.  
DR PRINTS; PR00791; PEPTIDASEA.  
DR PRODOM; PD004184; Peptidase\_M2; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOMN.1.  
SQ SEQUENCE 805 AA; 92491 MW; D2AAB4C2708EBE72 CRC64;

Query Match 99.9%; Score 4288; DB 4; Length 805;  
Best Local Similarity 99.9%; Pred. No. 5.8e-305;  
Matches 804; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSSSSWLLSLVAVTAOSTIEQAKTELDKFNHEADLFYQSSLASMNNTNITEENVQ 60
   1 MSSSSWLLSLVAVTAOSTIEQAKTELDKFNHEADLFYQSSLASMNNTNITEENVQ 60
Db 1 MSSSSWLLSLVAVTAOSTIEQAKTELDKFNHEADLFYQSSLASMNNTNITEENVQ 60
QY 61 MNNAAGDKWSAFLEKOSTLAQMYPLQEIQNLTVKLOLQALQNGSSVLSDESKRLNTIL 120
   61 MNNAAGDKWSAFLEKOSTLAQMYPLQEIQNLTVKLOLQALQNGSSVLSDESKRLNTIL 120
Db 61 MNNAAGDKWSAFLEKOSTLAQMYPLQEIQNLTVKLOLQALQNGSSVLSDESKRLNTIL 120
QY 121 NTMSTIYTGKVCNPDNPOECLELLEPGNEIMANSLDNERLMAWSEKSEVKGOLRPLY 180
   121 NTMSTIYTGKVCNPDNPOECLELLEPGNEIMANSLDNERLMAWSEKSEVKGOLRPLY 180
Db 121 NTMSTIYTGKVCNPDNPOECLELLEPGNEIMANSLDNERLMAWSEKSEVKGOLRPLY 180
QY 181 EEEVVLKNEMARANHYEDYGDYWRGDYEVNGVDYDYSRGOLIEDVHTFEETKPLYEHL 240
   181 EEEVVLKNEMARANHYEDYGDYWRGDYEVNGVDYDYSRGOLIEDVHTFEETKPLYEHL 240
Db 181 EEEVVLKNEMARANHYEDYGDYWRGDYEVNGVDYDYSRGOLIEDVHTFEETKPLYEHL 240
QY 241 HAYYRAKLMANAYPSYISPIGCLPAHLIDMGGRFNTLYSLTVPGQKPNIDYTDAMVQ 300
   241 HAYYRAKLMANAYPSYISPIGCLPAHLIDMGGRFNTLYSLTVPGQKPNIDYTDAMVQ 300
Db 241 HAYYRAKLMANAYPSYISPIGCLPAHLIDMGGRFNTLYSLTVPGQKPNIDYTDAMVQ 300
QY 301 AMDAORIFKEAEKFFVSVGLPNNTOGFENSMULTDPGNVOKAVCHPTAMDLGKDFRILM 360
   301 AMDAORIFKEAEKFFVSVGLPNNTOGFENSMULTDPGNVOKAVCHPTAMDLGKDFRILM 360
Db 301 AMDAORIFKEAEKFFVSVGLPNNTOGFENSMULTDPGNVOKAVCHPTAMDLGKDFRILM 360
QY 361 CTKVMTDDELTAHHEMGHIQYDMAVAAOPFLLRNGANGFHEAVGEINSLSAATPKHLKS 420
   361 CTKVMTDDELTAHHEMGHIQYDMAVAAOPFLLRNGANGFHEAVGEINSLSAATPKHLKS 420
Db 361 CTKVMTDDELTAHHEMGHIQYDMAVAAOPFLLRNGANGFHEAVGEINSLSAATPKHLKS 420
QY 421 IGLLSPDPOEDNETEINFLKQALTYGTLPTMYLEKRWNVFEGELIPKQOMKKWEM 480
   421 IGLLSPDPOEDNETEINFLKQALTYGTLPTMYLEKRWNVFEGELIPKQOMKKWEM 480
Db 421 IGLLSPDPOEDNETEINFLKQALTYGTLPTMYLEKRWNVFEGELIPKQOMKKWEM 480
QY 481 KREIVGVEPVPHDETYCDPASLFHVSNDYSFTRYTRTLVYOFQOALCOAAKHEGPLH 540
   481 KREIVGVEPVPHDETYCDPASLFHVSNDYSFTRYTRTLVYOFQOALCOAAKHEGPLH 540
Db 481 KREIVGVEPVPHDETYCDPASLFHVSNDYSFTRYTRTLVYOFQOALCOAAKHEGPLH 540
QY 541 KCDISNSTEAGQKLFNNMLRLGKSEPTLALENVGAKNNVRLNTEFEPLFTWLKDQNK 600
   541 KCDISNSTEAGQKLFNNMLRLGKSEPTLALENVGAKNNVRLNTEFEPLFTWLKDQNK 600
Db 541 KCDISNSTEAGQKLFNNMLRLGKSEPTLALENVGAKNNVRLNTEFEPLFTWLKDQNK 600
QY 601 NSFVGSTDMSPYADOSIKVRISLSKSGADKAYEMNDNEMYLERSSVAAKROYFLKAVN 660
   601 NSFVGSTDMSPYADOSIKVRISLSKSGADKAYEMNDNEMYLERSSVAAKROYFLKAVN 660
Db 601 NSFVGSTDMSPYADOSIKVRISLSKSGADKAYEMNDNEMYLERSSVAAKROYFLKAVN 660
QY 661 QMILFGEEDVRANLKPRISFNEFVTAAPKNVSDIIPTEVEKAIKMSRSRIINDAFLRNDN 720
   661 QMILFGEEDVRANLKPRISFNEFVTAAPKNVSDIIPTEVEKAIKMSRSRIINDAFLRNDN 720

```

Db 661 QMIFGEEDVAVANLKPRISFNEFVTAPKRVSDIIPREVEKAIKMSRSRINDAFRLDN 720  
 QY 721 SLEFLGIQPTLGPNNOPPVSTIMLIVFGVYGVIVGIVLITFGIDRKKKKRARGSENP 780  
 Db 721 SLEFLGIQPTLGPNNOPPVSTIMLIVFGVYGVIVGIVLITFGIDRKKKKRARGSENP 780  
 QY 761 YASIDISKGENNPGFQNTDVTQTSF 805  
 Db 761 YASIDISKGENNPGFQNTDVTQTSF 805

## RESULT 3

Q9UFZ6 PRELIMINARY: PRT: 804 AA.  
 ID Q9UFZ6  
 AC 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Hypothetical 92.3 kDa protein (Fragment).  
 GN DKEP34A014.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OC NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TESTIS;  
 RA Mambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL110224; CAB53682.1; -  
 DR MEROPS: M02\_006; -  
 DR InterPro: IPR001548; Peptidase\_M2.  
 DR InterPro: IPR001130; Zn\_MTPeptidase.  
 DR Pfam: PF01401; Peptidase\_M2; 1.  
 DR PRINTS: PR007191; PEPTIDPTASEA.  
 DR ProDom: PD004184; Peptidase\_M2; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 804 AA; 92340 MW; 91FF391074CB5DA9 CRC64;

Query Match 99.8%; Score 4281; DB 4; Length 804;  
 Best Local Similarity 99.9%; Pred. No. 1.9e-304;  
 Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSSSWLLSLVAVTAAGSTIEQAKTFLDKFNHEAEDLFYQSSLASWNTNTITEENVQ 61  
 Db 1 SSSSWLLSLVAVTAAGSTIEQAKTFLDKFNHEAEDLFYQSSLASWNTNTITEENVQ 60  
 QY 62 MNNAAGKMSAFLEKQSTLAQMTPLQEIQNTLVKLOLALQONGSSVSEDSKRLNTLN 121  
 Db 61 MNNAAGKMSAFLEKQSTLAQMTPLQEIQNTLVKLOLALQONGSSVSEDSKRLNTLN 120  
 QY 122 TNSITVSTGKVCNPNPQECILLPEGLNEMANSLDYNERLWAMESRSEYKQLRPLYE 181  
 Db 121 TNSITVSTGKVCNPNPQECILLPEGLNEMANSLDYNERLWAMESRSEYKQLRPLYE 180  
 QY 182 EYVVLKEMARAHNYEDYGDYRGDYEVNGVDGYDYSRQQLLEDVHEHFEELKPLYEHL 241  
 Db 181 EYVVLKEMARAHNYEDYGDYRGDYEVNGVDGYDYSRQQLLEDVHEHFEELKPLYEHL 240  
 QY 242 AYVRAKLMAAYPSYISPTGCLPAHLGLDMGGRFMTNLVSLVYFGQKPNIDVTDAWVQA 301  
 Db 241 AYVRAKLMAAYPSYISPTGCLPAHLGLDMGGRFMTNLVSLVYFGQKPNIDVTDAWVQA 300  
 QY 302 WDAORITFEAKFEFVSGLPNTQGFENSMITLDPGNVQKAVCHPAMPDLGKDFRILMC 361  
 Db 301 WDAORITFEAKFEFVSGLPNTQGFENSMITLDPGNVQKAVCHPAMPDLGKDFRILMC 360  
 QY 362 TKVIMDDPLTAHHEMGHLYQYDMAVYAAOPFLLRNGANGFHEAVGEIMSLSAATPKLKST 421  
 Db 361 TKVIMDDPLTAHHEMGHLYQYDMAVYAAOPFLLRNGANGFHEAVGEIMSLSAATPKLKST 420

QY 422 GLSPDFQEDNETEINFLKQALTIYGTLPFTYMLFKRMNVEKGEIPKQDMKKMKWK 481  
 Db 421 GLSPDFQEDNETEINFLKQALTIYGTLPFTYMLFKRMNVEKGEIPKQDMKKMKWK 480  
 QY 482 REIVGVAPPHDFTYCDPASPFLFVSNDSYFIERYTETLYQFOFQALCOAAKHGEPHLK 541  
 Db 481 REIVGVAPPHDFTYCDPASPFLFVSNDSYFIERYTETLYQFOFQALCOAAKHGEPHLK 540  
 QY 542 CDISNSTEAGCKLFNMLRLGKSEPTLALENVCAKMANVRLPLNFEPLFTMLKQONKN 601  
 Db 541 CDISNSTEAGCKLFNMLRLGKSEPTLALENVCAKMANVRLPLNFEPLFTMLKQONKN 600  
 QY 602 SFVGNSTQMSPEYADOSTIVRISIKSALGDKAYENNDENMLFRSSVAYAMROYFLKYNQ 661  
 Db 601 SFVGNSTQMSPEYADOSTIVRISIKSALGDKAYENNDENMLFRSSVAYAMROYFLKYNQ 660  
 QY 662 MIFGEEDVAVANLKPRISFNEFVTAPKRVSDIIPREVEKAIKMSRSRINDAFRLDN 721  
 Db 661 MIFGEEDVAVANLKPRISFNEFVTAPKRVSDIIPREVEKAIKMSRSRINDAFRLDN 720  
 QY 722 LEFLGIQPTLGPNNOPPVSTIMLIVFGVYGVIVGIVLITFGIDRKKKKRARGSENP 781  
 Db 721 LEFLGIQPTLGPNNOPPVSTIMLIVFGVYGVIVGIVLITFGIDRKKKKRARGSENP 780  
 QY 782 ASIDISKGENNPGFQNTDVTQTSF 805  
 Db 781 ASIDISKGENNPGFQNTDVTQTSF 804

## RESULT 4

Q8R010 PRELIMINARY: PRT: 805 AA.  
 ID Q8R010  
 AC 01-JUN-2002 (Tremblrel. 21, Created)  
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE RIKEN cDNA 2010305L05 gene.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBL\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=KIDNEY;  
 RA Strausberg R.;  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.  
 SQ EMBL: BC026801; AAH26801.1; -  
 SQ SEQUENCE 805 AA; 92367 MW; D8B883AAC966A8D9 CRC64;

Query Match 83.4%; Score 3579; DB 11; Length 805;  
 Best Local Similarity 82.1%; Pred. No. 4.2e-253;  
 Matches 661; Conservative 60; Mismatches 84; Indels 0; Gaps 0;

QY 1 MSSSWLLSLVAVTAAGSTIEQAKTFLDKFNHEAEDLFYQSSLASWNTNTITEENVQ 60  
 Db 1 MSSSWLLSLVAVTAAGSTIEQAKTFLDKFNHEAEDLFYQSSLASWNTNTITEENVQ 60  
 QY 61 MNNAAGKMSAFLEKQSTLAQMTPLQEIQNTLVKLOLALQONGSSVSEDSKRLNTLN 120  
 Db 61 MNNAAGKMSAFLEKQSTLAQMTPLQEIQNTLVKLOLALQONGSSVSEDSKRLNTLN 120  
 QY 121 TNSITVSTGKVCNPNPQECILLPEGLNEMANSLDYNERLWAMESRSEYKQLRPLYE 180  
 Db 121 TNSITVSTGKVCNPNPQECILLPEGLNEMANSLDYNERLWAMESRSEYKQLRPLYE 180  
 QY 181 EYVVLKEMARAHNYEDYGDYRGDYEVNGVDGYDYSRQQLLEDVHEHFEELKPLYEHL 240  
 Db 181 EYVVLKEMARAHNYEDYGDYRGDYEVNGVDGYDYSRQQLLEDVHEHFEELKPLYEHL 240  
 QY 241 AYVRAKLMAAYPSYISPTGCLPAHLGLDMGGRFMTNLVSLVYFGQKPNIDVTDAWVQA 300  
 Db 241 AYVRAKLMAAYPSYISPTGCLPAHLGLDMGGRFMTNLVSLVYFGQKPNIDVTDAWVQA 300

RESULT	5
099N71	
ID	099N71
AC	099N71;
DT	01-JUN-2001 (TREMBLrel. 17, Preliminary; PRt, 798 AA.
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DN	Angiotensin-converting enzyme-related carboxypeptidase.
GN	Angiotensin-converting enzyme-related carboxypeptidase.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxId=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Komatsu T., Sugano S., Imai J., Suzuki Y., Hanaoka K., Yanao Y.,
RA	Hida M., Taniguchi A., Muroi S.;
RT	Molecular cloning, mRNA expression, and chromosomal localization of
RT	mouse Angiotensin-converting Enzyme-related Carboxypeptidase.;
RT	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
EMBL	EMBL: AB033181. BAB40431.1; -.
MEMOPS	MEMOPS: M02.006.

Query Match	Best Local Similarity	Score 3485; DB 11; Length 798;
MSD; MGI:1917256; 2010305L05R1K.	81.2%;	
InterPro; IPR001548; Peptidase_M2.	82.2%;	
InterPro; IPR001680; WD40.		
InterPro; IPR000130; Zn_MTPeptase.		
Pfam; PF01401; Peptidase_M2; 1.		
PRINTS; PR00791; PeptIDPTASEA.		
PRODOM; PD004184; Peptidase_M2; 1.		
PROSITE; PS00678; WD_REPEAT_1; UNKNOWN_1.		
PROSITE; PS00142; ZINC_PROTASE; UNKNOWN_1.		
Carboxypeptidase.		
SEQUENCE 798 AA; 91943 MW; 403AEA29D55725A4 CRC64;		

099N70	PRELIMINARY;	PRT;	353 AA.
ID	099N70		
AC	099N70;		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Angiogenin-converting enzyme-related carboxypeptidase.		
GN	ACE2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		



RP SEQUENCE FROM N.A.  
 RA Komatsu T., Sugano S., Imai J., Suzuki Y., Hanaoka K., Yamada Y.,  
 RA Hida M., Tanigami A., Murai S.,  
 RT "Molecular cloning, mRNA expression, and chromosomal localization of  
 RT mouse Angiotensin-converting Enzyme-Related Carboxypeptidase-";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Komatsu T., Sugano S., Suzuki Y., Hanaoka K., Yamada Y.,  
 RT "Molecular cloning of ACE2",  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB053182; BAB40432.1; -  
 DR MEROPS: M02.006; -  
 DR Interpro: IPR001548; Peptidase\_M2.  
 DR Interpro: IPR001680; WD40.  
 DR Pfam: PF01401; Peptidase\_M2; 1.  
 DR PRINTS: PR00791; PEPTIDPASEA.  
 DR PROSITE: PS00678; WD\_REPEATS\_1; UNKNOWN\_1.  
 DR Carboxypeptidase.  
 SQ SEQUENCE 353 AA; 40442 MW; D17B71141EB4AF5B CRC64;

Query Match 35.9%; Score 1539; DB 11; Length 353;  
 Best Local Similarity 81.2%; Pred. No. 2.1e-104;  
 Matches 285; Conservative 22; Mismatches 44; Indels 0; Gaps 0;

QY 1 MSSSSMLLSLVAVYTAOSTIEQAKTFLDKFNHEDLYQSSLASMNNTNTEENQ 60  
 DB 1 MSSSSMLLSLVAVYTAOSTIEQAKTFLDKFNHEDLYQSSLASMNNTNTEENQ 60  
 QY 61 MNMNGDKMSAFLEKQSTLAOMYPLQETONTLVKLOALQONGSSVLSDEKSKRLNTIL 120  
 DB 61 MNMNGDKMSAFLEKQSTLAOMYPLQETONTLVKLOALQONGSSVLSDEKSKRLNTIL 120  
 QY 121 NMSTYSTGKVCNPDNPOECLELLEPGNEIMANSIDYERLWAMESRSEVQKRLPLY 180  
 DB 121 NMSTYSTGKVCNPDNPOECLELLEPGNEIMANSIDYERLWAMESRSEVQKRLPLY 180  
 QY 121 NMSTYSTGKVCNPDNPOECLELLEPGNEIMANSIDYERLWAMESRSEVQKRLPLY 180  
 DB 121 NMSTYSTGKVCNPDNPOECLELLEPGNEIMANSIDYERLWAMESRSEVQKRLPLY 180  
 QY 181 EEEVVLKNNMARNHEDYGDYWRGDYVNGVDGYDSRGQLIEDVEHFEIKPLYEHL 240  
 DB 181 EEEVVLKNNMARNHEDYGDYWRGDYVNGVDGYDSRGQLIEDVEHFEIKPLYEHL 240  
 QY 181 EEEVVLKNNMARNHEDYGDYWRGDYVNGVDGYDSRGQLIEDVEHFEIKPLYEHL 240  
 DB 181 EEEVVLKNNMARNHEDYGDYWRGDYVNGVDGYDSRGQLIEDVEHFEIKPLYEHL 240  
 QY 241 HAYVRKLNMAVPSYSPICGLPAHLIDGMGRFNTNLSLTPVGQKPNIDVTDAVQ 300  
 DB 241 HAYVRKLNMAVPSYSPICGLPAHLIDGMGRFNTNLSLTPVGQKPNIDVTDAVQ 300  
 QY 301 AMDQRIFFEAKEFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 351  
 DB 301 AMDQRIFFEAKEFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 351  
 QY 301 GMDAERIFQEAKEFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 351  
 DB 301 GMDAERIFQEAKEFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 351

RESULT 7  
 Q9GLN6 PRELIMINARY; PRT: 732 AA.  
 ID Q9GLN6  
 AC Q9GLN6;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE Dipeptidyl carboxy peptidase 1 testicular form.  
 GN DCP1.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20469400; PubMed=11013071;  
 RA Dufour C., Casane D., Denton D., Wickings J., Corvol P.,  
 RA Jeunemaitre X.,  
 RT "Human-Chimpanzee DNA sequence variation in the four major genes of  
 RT the renin-angiotensin system."  
 RL Genomics 69:14-26(2000)  
 DR EMBL: AF193473; AAC31359.1; JOINED.  
 DR EMBL: AF193473; AAC31359.1; JOINED.

DR EMBL: AF193474; AAC31359.1; JOINED.  
 DR EMBL: AF193475; AAC31359.1; JOINED.  
 DR EMBL: AF193476; AAC31359.1; JOINED.  
 DR EMBL: AF193477; AAC31359.1; JOINED.  
 DR EMBL: AF193478; AAC31359.1; JOINED.  
 DR EMBL: AF193479; AAC31359.1; JOINED.  
 DR EMBL: AF193480; AAC31359.1; JOINED.  
 DR EMBL: AF193481; AAC31359.1; JOINED.  
 DR EMBL: AF193482; AAC31359.1; JOINED.  
 DR EMBL: AF193483; AAC31359.1; JOINED.  
 DR EMBL: AF193484; AAC31359.1; JOINED.  
 DR EMBL: AF193485; AAC31359.1; JOINED.  
 DR Interpro: IPR001548; Peptidase\_M2.  
 DR Interpro: IPR00130; Zn\_MTpeptidase.  
 DR Pfam: PF01401; Peptidase\_M2; 1.  
 DR PRINTS: PR00791; PEPTIDPASEA.  
 DR PRODOM: PD004184; Peptidase\_M2; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 SQ SEQUENCE 732 AA; 83428 MW; C4306443A47E74B CRC64;

Query Match 31.1%; Score 1336; DB 6; Length 732;  
 Best Local Similarity 41.6%; Pred. No. 4.3e-89;  
 Matches 258; Conservative 119; Mismatches 205; Indels 38; Gaps 10;

QY 15 TAAOS---TTEQAKTFLDKFNHEDLYQSSLASMNNTNTEB-----NWOM 62  
 DB 15 TAAOS---TTEQAKTFLDKFNHEDLYQSSLASMNNTNTEB-----NWOM 62  
 QY 63 NNAGDKMSAFLEKQSTLAOMYPLQETONTLVKLOALQONGSSVLSDEKSKRLNTIL 122  
 DB 63 NNAGDKMSAFLEKQSTLAOMYPLQETONTLVKLOALQONGSSVLSDEKSKRLNTIL 122  
 QY 121 NHT-----LKGTQARRFDVNOONTTIRIKKVOEDERAPLAQEEENKILTD 172  
 DB 121 NHT-----LKGTQARRFDVNOONTTIRIKKVOEDERAPLAQEEENKILTD 172  
 QY 123 MSTYSTGKVCNPDNPOECLELLEPGNEIMANSIDYERLWAMESRSEVQKRLPLYEHL 182  
 DB 123 MSTYSTGKVCNPDNPOECLELLEPGNEIMANSIDYERLWAMESRSEVQKRLPLYEHL 182  
 QY 173 METTYSVATVCTNG--SCLQLEPDLTVNMAVTSKRYEDLLMAGWDKGRALQIFPK 230  
 DB 173 METTYSVATVCTNG--SCLQLEPDLTVNMAVTSKRYEDLLMAGWDKGRALQIFPK 230  
 QY 183 YVVLKNNMARNHEDYGDYWRGDYVNGVDGYDSRGQLIEDVEHFEIKPLYEHL 242  
 DB 183 YVVLKNNMARNHEDYGDYWRGDYVNGVDGYDSRGQLIEDVEHFEIKPLYEHL 242  
 QY 231 YVELINQARLNGYVDAGSMRSMTETPSLE-----QDLERLFQELQPLYLHA 280  
 DB 231 YVELINQARLNGYVDAGSMRSMTETPSLE-----QDLERLFQELQPLYLHA 280  
 QY 243 YVRAKLNMAVPSYSPICGLPAHLIDGMGRFNTNLSLTPVGQKPNIDVTDAVQ 301  
 DB 243 YVRAKLNMAVPSYSPICGLPAHLIDGMGRFNTNLSLTPVGQKPNIDVTDAVQ 301  
 QY 281 YVRAKLNMAVPSYSPICGLPAHLIDGMGRFNTNLSLTPVGQKPNIDVTDAVQ 340  
 DB 281 YVRAKLNMAVPSYSPICGLPAHLIDGMGRFNTNLSLTPVGQKPNIDVTDAVQ 340  
 QY 302 WDAQRIFFEAKEFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 360  
 DB 302 WDAQRIFFEAKEFEVSVGLPMTQGFENSMULTDPGNQKAVCHPTAMD 360  
 QY 341 WTPRRMFKKADDFETSLGILPVPPEFWNKSMEKFTDGEVYCHASAMDFNGDFRQ 400  
 DB 341 WTPRRMFKKADDFETSLGILPVPPEFWNKSMEKFTDGEVYCHASAMDFNGDFRQ 400  
 QY 361 CTVTNDDFLTAHENGHIQYDMAVAAOPFLIRNGANEGFHEAVGEISLSAATPKILKS 420  
 DB 361 CTVTNDDFLTAHENGHIQYDMAVAAOPFLIRNGANEGFHEAVGEISLSAATPKILKS 420  
 QY 401 CTVTNDDFLTAHENGHIQYDMAVAAOPFLIRNGANEGFHEAVGEISLSAATPKILKS 460  
 DB 401 CTVTNDDFLTAHENGHIQYDMAVAAOPFLIRNGANEGFHEAVGEISLSAATPKILKS 460  
 QY 421 IGLSPDFQDNTEINFLIKQALTVGLPTLYMEKRRMVFEGELPKQOMKKNWEM 480  
 DB 421 IGLSPDFQDNTEINFLIKQALTVGLPTLYMEKRRMVFEGELPKQOMKKNWEM 480  
 QY 461 LNLISSEGSQD--PHDINFLMKALDKIAFIPSYLDVQRMVFEQSTKEKYNQDNEMSL 519  
 DB 461 LNLISSEGSQD--PHDINFLMKALDKIAFIPSYLDVQRMVFEQSTKEKYNQDNEMSL 519  
 QY 481 KREIVGVEVEPPDEYCDPASLFFHSNDYSFIRYTRFLYQFOEALCAQAHEGFLH 540  
 DB 481 KREIVGVEVEPPDEYCDPASLFFHSNDYSFIRYTRFLYQFOEALCAQAHEGFLH 540  
 QY 520 RLKYGGLCPVPRTQGDPEGAKFHPSSVPIRIVVSIIIOFQHEALCAQAHEGFLH 579  
 DB 520 RLKYGGLCPVPRTQGDPEGAKFHPSSVPIRIVVSIIIOFQHEALCAQAHEGFLH 579  
 QY 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVGAKNNVRPLPPELFTWLKDKQ 600  
 DB 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVGAKNNVRPLPPELFTWLKDKQ 600  
 QY 580 KCDIYQKSGKQGRATAMKLGFSRPMPEAMOLITGPNMSASAMLYFRLDMLWLTENE 639  
 DB 580 KCDIYQKSGKQGRATAMKLGFSRPMPEAMOLITGPNMSASAMLYFRLDMLWLTENE 639  
 QY 601 --NSFVGW--STDWSPYADQS 617  
 DB 601 --NSFVGW--STDWSPYADQS 617  
 QY 640 LHCEKLGMPQYMTNPNSARS 659  
 DB 640 LHCEKLGMPQYMTNPNSARS 659

RESULT 8  
 Q9GLN7 PRELIMINARY; PRT: 1304 AA.  
 ID Q9GLN7  
 AC Q9GLN7;

```

DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Dipeptidyl carboxy peptidase 1.
GN DCPL1
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_Taxid=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20469400; PubMed=11013071;
RA Dufour C., Casane D., Denton D., Wickings J., Corvol P.,
RA Jeunemaitre X.;
RT "Human-Chimpanzee DNA sequence variation in the four major genes of
RT the renin angiotensin system.";
RL Genomics 69:14-26(2000)
DR EMBL; AF193466; AAC31358.1; -
DR EMBL; AF193462; AAC31358.1; JOINED.
DR EMBL; AF193463; AAC31358.1; JOINED.
DR EMBL; AF193464; AAC31358.1; JOINED.
DR EMBL; AF193465; AAC31358.1; JOINED.
DR EMBL; AF193466; AAC31358.1; JOINED.
DR EMBL; AF193467; AAC31358.1; JOINED.
DR EMBL; AF193468; AAC31358.1; JOINED.
DR EMBL; AF193469; AAC31358.1; JOINED.
DR EMBL; AF193470; AAC31358.1; JOINED.
DR EMBL; AF193471; AAC31358.1; JOINED.
DR EMBL; AF193472; AAC31358.1; JOINED.
DR EMBL; AF193473; AAC31358.1; JOINED.
DR EMBL; AF193474; AAC31358.1; JOINED.
DR EMBL; AF193475; AAC31358.1; JOINED.
DR EMBL; AF193476; AAC31358.1; JOINED.
DR EMBL; AF193477; AAC31358.1; JOINED.
DR EMBL; AF193478; AAC31358.1; JOINED.
DR EMBL; AF193479; AAC31358.1; JOINED.
DR EMBL; AF193480; AAC31358.1; JOINED.
DR EMBL; AF193481; AAC31358.1; JOINED.
DR EMBL; AF193482; AAC31358.1; JOINED.
DR EMBL; AF193483; AAC31358.1; JOINED.
DR EMBL; AF193484; AAC31358.1; JOINED.
DR EMBL; AF193485; AAC31358.1; JOINED.
DR MEROPS; M02.001; -.
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR000130; Zn_MTEpeptidse.
DR Pfam; PF01401; Peptidase_M2; 2.
DR PRINTS; PR00791; PEPDIPASE.
DR ProDom; PD004184; Peptidase_M2; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN.1.
SQ SEQUENCE 1304 AA; 149369 MW; DCF728D0BA0F1314 CRC64;

Query Match 31.0%; Score 1329; DB 6; Length 1304;
Best Local Similarity 41.6%; Pred. No. 3.3e-88;
Matches 254; Conservative 118; Mismatches 205; Indels 34; Gaps 9;

QY 20 TIEQAKTFLDKFNHEADLFYQSSLASWNTNTTTE-----VQNNNNAGDWSA 71
DB 642 TDEAEKSFVEEYDRTSOVWNEVEAMWNTNTTTSKILLQKNQJANH----- 695
QY 72 FLKEGSLAOMVPLDEIONLVKLOLQALQONGSSVSEDSKRLNTLNTMSTYSTGK 131
DB 696 --LKGTOARRFDVNOJLONTTIRIKKVKQDLERAAALPAQLEBYNKILLDMETTSVAT 753
QY 132 VCNPDNPECLLEPGLMEIMANSIDYNERLWAMESRSVEGKQRLPYLEYEYVVKEMA 191
DB 754 VCHTNG--SCQLLEPDLINWATSRKYEEDLWAMEGWRDAGRAILLOFYERYVELNOAA 811
QY 192 RANHEEDYGVYRGDYEYVNGVDYDSRGQIEDVEHTEFEIKPLYEHLHAAYRAKLMA 251
DB 812 RLVNGVDAGDSWRSRYETPSLE-----QDLERLFEQLDPLVYNLHAAYRRALHNR 861
QY 252 Y-PSYISPIGCLPAHLGLGDMWGRFWNTLVSLTYVFGKRNIDVTADAVQDADQRIEKE 310

```

```

DB 862 YGQNHINLEGRLPAHLGMMWQTSNITLYDLVPPPSAPSDTTEAMLKQGWPRRMFKE 921
QY 311 AEKFPYVGLPMMTOGFENSMILTDPGVQKAVCPHTAMDGKG-DFRIIMCTKVWDDF 369
DB 922 ADDEFTSLGLLVPPPEFNKRSMLKEKPTDGRVYVCHASAMDFYDNCKDPRINQCTVNLDEL 981
QY 370 LTAHHEMGIQYDMAVYAAQPLFRNGAGNEHVEGELMSLSATPRNHLKSTGLSPDFQ 429
DB 982 VVAHHEMGIQYFMQYKPLPALRGAGNPGFHEALIGVALSVTPKHLNLSLSEGG 1041
QY 430 EDNETEINFLKQALTTVGLPFTYMLEKWMVFEKEIPKDDMKKWMKREIYGVYE 489
DB 1042 SD-EHDINFLMKALDKTAFIPFSYLVQDMWRVFDSDIIRKENYNDPMWLSRLKYGLCP 1100
QY 490 PYPHEETCDPASLPHVSNDSFTRYTRILYQFOEALCQAKHGGPLMKCDISNSTE 549
DB 1101 PVPRTQGFDPGAKFHIPSSVPIRYFVSFIQFOHEALCQAAHGTGPLHKCDIYQSK 1160
QY 550 AGOKLFNMLRLKSEPTLALENVYCAKNNVRLNTPPELFTWLKDQNK--NSVGV- 606
DB 1161 AGORLATAMKLGFSRPPPEAMQILITGPNNASAMLSYFPLDMLRTNELHGEKLGWP 1220
QY 607 STDWSPRADOS 617
DB 1221 QYNWTPNSARS 1231

RESULT 9
Q9EQW9 PRELIMINARY; PRT; 1313 AA.
AC Q9EQW9
ID Q9EQW9
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Angiotensin-converting enzyme.
GN ACE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F344/N; TISSUE=LUNG;
RA Jafarian-Thrani M., Listwak S., Barrientos R.M., Michaud A.,
RA Corvol P., Sternberg E.M.;
RT "Characterization of a missense mutation in the angiotensin I-
RT converting enzyme cDNA in exudative inflammation resistant F344/N
RT rats.";
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF201331; AAC35596.1; -
DR InterPro; IPR001548; Peptidase_M2.
DR InterPro; IPR000130; Zn_MTEpeptidse.
DR Pfam; PF01401; Peptidase_M2; 2.
DR PRINTS; PR00791; PEPDIPASE.
DR ProDom; PD004184; Peptidase_M2; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN.1.
SQ SEQUENCE 1313 AA; 150941 MW; E888341562542B04 CRC64;

Query Match 30.5%; Score 1310; DB 11; Length 1313;
Best Local Similarity 42.0%; Pred. No. 8.1e-87;
Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps 7;

QY 20 TIEQAKTFLDKFNHEADLFYQSSLASWNTNTTTE-----VQNNNNAGDWSA 79
DB 650 TDEKANKNFVEEYDRTAKVLNNEVEAMWNTNTTTSKILLQKNKESNHTLKYGTW 709
QY 80 AOMVPLDEIONLVKLOLQALQONGSSVSEDSKRLNTLNTMSTYSTGVCVNDPQ 139
DB 710 AKTFDVSNFQNSTIKRIKKVQVNDRAVLPNPELEBYNQILLDMETTSVANVCYTG-- 767
QY 140 ECLLEPGLMEIMANSIDYNERLWAMESRSVEGKQRLPYLEYEYVVKEMARANHYEDY 199

```

Db 768 TCLSLEPDLTNIMATSKRYEELLWMSKRDVKGRAILPEFPKYVDPSNKIKANGYSOA 827  
 Qy 200 GDIYRGDIYVNGVDGYYSRGQLEIYEHFEETIKPLIYELHAYVRAKIMNVPY-YSIP 258  
 Db 828 GDSWRSSSESDLE-----QDLEKYLQELQPLYLHAYVRSLSHRHAGSEYIM 877  
 Qy 259 IGCPLPAHLGDMGREFWNTLSLTPFGOKPNIDVDAMDQAMDQARTKEAEKEFPVS 318  
 Db 878 DGPFAHLGDMGREFWNTLSLTPFGOKPNIDVDAMDQAMDQARTKEAEKEFPVS 937  
 Qy 319 GLPNTQGFENSMITDGNVQKAVCHPTAMDLSG-DEFILMCTKYMTDFLTAHENG 377  
 Db 938 GLPVPPEFWMKSMLEKPTDREYVCHASAMDFYNGDFRIKOCSTYVMELEYAHENG 997  
 Qy 378 HIOYMAAQAOPFLLRNANGFHEAVGEIMSLAATPKHLKSGILSPDQEDNETEIN 437  
 Db 998 HIOYMOYKOLPYTFREANGFHEALIGDVALSVTPKHLHSLNLISSE-GSGYEHDIN 1056  
 Qy 438 FLKCALITVGTLPFTYMLERKRWKVEKGETPKDQMKKWKEMKREIVGVPEPHDEY 497  
 Db 1057 FLKMAALDKIAFIPPSYLDQMRWVFDGSIKENYQENWMSLTKYQGLCPVPSQGD 1116  
 Qy 498 CDPALFHSNDYSFIRYRTYTRLYOFQFOALCOAKHEGPLKCDISNSTEAGOKLFM 557  
 Db 1117 FPGSKFHPANVPYIRKIFISITIOFQHEALCRAAGHTGPLYKCDIYOSKEAGKLLADA 1176  
 Qy 558 LRLGSEPTLALENVGAKNNVRLNTEPLETTLKDONK--NSFVGW-STDWSP 612  
 Db 1177 MKLGYSKQPEAKMKTITGQPNASASAIMYKPLETTEMLVTENRRHGETLGWPEYTW 1234

## RESULT 10

Q15540 PRELIMINARY: PRG: 694 AA.  
 ID Q15540: AC Q15540: 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)  
 DE DCP1 protein.  
 GN DCP1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90046671; PubMed=2554286;  
 RA Ehlers M.R., Fox E.A., Strydom D.J., RJordan J.F.;  
 RT "Molecular cloning of human testicular angiotensin-converting enzyme:  
 RT the testis isozyme is identical to the C-terminal half of endothelial  
 RT angiotensin-converting enzyme.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7741-7745(1989).  
 DR EMBL: M26658; AAA60612.1;  
 DR InterPro: IPR001548; Peptidase\_M2.  
 DR InterPro: IPR00130; Zn\_MTpeptidase.  
 DR Pfam: PF01401; Peptidase\_M2; 1.  
 DR PRINTS: PR00791; PEPDIPITASEA.  
 DR PRODOM: PD004184; Peptidase\_M2; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN.1.  
 DR SEQUENCE 694 AA; 79333 MW; 57C0F9C5A5BCA119 CRC64;

Query Match 30.4%; Score 1303; DB 4; Length 694;  
 Best Local Similarity 42.7%; Pred. No. 1e-86;  
 Matches 250; Conservative 107; Mismatches 195; Indels 34; Gaps 9;

Qy 45 LASWNTNITEE-----NCONNNAGDKMSAFLEKSTIAQMPLOEITONLVKQ 96  
 Db 57 IATKYNNTTETSKILQKNOQIANHT-----LKYGTAQKRDVNOJONTTICKR 108  
 Qy 97 LOALQONSSVLSDEKSKRLNTIIMSTIYSGVCNPDPGCLLEGLNEIMANSU 156  
 Db 109 IKVODLERALPAQELIEYKILLDMETTSVATVCHPNG--SCDLEPDLINMAISR 166

Qy 157 DYNRLAMESNBSNGVKOLRPLYEEVYLKNEMARANHEDYGDYRGDYEVEVNGVDYD 216  
 Db 167 KYEDLWAMEBGRDKARAILQFPKVELINQAAHLNGYVADGDSRSMETPSLE--- 223  
 Qy 217 YSRQGLIYEHFEETIKPLIYELHAYVRAKIMNAV-PSYISIPGCLPAHLGDMGREF 275  
 Db 224 -----QDLEKYLQELQPLYLHAYVRSLSHRHAGSEYIM 276  
 Qy 276 TNLVSLTPFGOKPNIDVDAMDQAMDQARTKEAEKEFPVSGLPNTQGFENSMITD 335  
 Db 277 SNIVDLVVPFSPASMDTEAMKQGTVPKREADEFTSLGLLPVPEPFWKSMLEK 336  
 Qy 336 PGNVQKAVCHPTAMDLSG-DEFILMCTKYMTDFLTAHENGHIQYMAAQAOPFLRN 394  
 Db 337 PLDGEVYCHASAMDFYNGDFRIKOCSTYVMELEYAHENGHIQYMOYKOLPYALRE 396  
 Qy 395 GANEGFHEAVGEIMSLAATPKHLKSGILSPDQEDNETEINFLKCALITVGTLPFTY 454  
 Db 397 GANPFHEALIGDVALSVTPKHLHSLNLISSEGGSD-EHDINFLKMAALDKIAFIPFS 455  
 Qy 455 MEKRWVVEKGETPKDQMKKWKEMKREIVGVPEPHDEYCDPALSLEHVSNDYSFIR 514  
 Db 456 LVDQMRWVFDGSIKENYQENWMSLTKYQGLCPVPSQGD 515  
 Qy 515 YTRTLYOFQFOALCOAKHEGPLKCDISNSTEAGOKLFMLRLGSEPTLALENV 574  
 Db 516 YFVSFLIQFOFHEALCOAGHTGPLYKCDIYOSKEAGRLATAKLGFSPWPEAMQLT 575  
 Qy 575 GAKNNVRLNTEPLETTLKDONK--NSFVGW-STDWSPYDAS 617  
 Db 576 GQPNASASAMLSYKPLDMLTENELEKLGWPOYWTNSARS 621

## RESULT 11

Q9NDS8 PRELIMINARY: PRG: 648 AA.  
 ID Q9NDS8: AC Q9NDS8: 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)  
 DE Bcdysteroind-inducible angiotensin-converting enzyme-related gene  
 DE product.  
 OS Bombyx mori (Silk moth).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
 OC Bombycoidea; Bombycidae; Bombyx.  
 OC NCBI\_TaxID=7091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=C108; TISSUE=WING IMAGINAL DISK;  
 RX MEDLINE=20556050; PubMed=11102839;  
 RA Quan G.X., Mita K., Okano K., Shimada T., Ugaia N., Xia Z., Goto N.,  
 RA Kanke E., Kawasaki H.;  
 RT "Isolation and expression of the ecdysteroid-inducible angiotensin-  
 RT converting enzyme-related gene in wing discs of Bombyx mori.";  
 RL Insect Biochem. Mol. Biol. 31:97-103(2001).  
 DR EMBL: AB026110; BAA97657.1;  
 DR MEROPS: M02.002;  
 DR InterPro: IPR001548; Peptidase\_M2.  
 DR InterPro: IPR00130; Zn\_MTpeptidase.  
 DR Pfam: PF01401; Peptidase\_M2; 1.  
 DR PRINTS: PR00791; PEPDIPITASEA.  
 DR PRODOM: PD004184; Peptidase\_M2; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN.1.  
 DR SEQUENCE 648 AA; 74917 MW; 9A740AA9FCACEBF0 CRC64;

Query Match 25.4%; Score 1089.5; DB 5; Length 648;  
 Best Local Similarity 35.6%; Pred. No. 3.9e-71;  
 Matches 232; Conservative 125; Mismatches 259; Indels 35; Gaps 11;

Qy 11 LVAVTAOSTIEQAK-TFLDKENHEADLFT-----QSLASWNTNITEEN 58  
 Db 10 LIAAIVAVFIVATGGRDPLEAREHAREYMLHLKATGLKRNRASLAEWETSNITKEN 69





```

[2]
SEQUENCE FROM N.A.
RP STRAIN-BERKELEY;
RA Ceiniker S.E., George R.A., Galle R., Svirska R.R., Hoskins R.A.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazet R.G., Chavez C., Chew M.,
RA Doyle C.M., Farfan D.E., Flanagan J., Houston K.A., Hummel S.R.,
RA Kaira K., Kearney L., Kim S.H., Lee B., Lomolan M.A., Mak J.,
RA Mada P., Mok M.S., Moshrefi A.R., Moshrefi M., Nixon R., Pacleb J.M.,
RA Park S., Pfeiffer B., Punch E., Snir E., Twomey B., Wan K.H.,
RA Whitelaw K.R., Yee A., Zhang R., Zieran L.L., Kimmel B.E.,
RL Submitted (MAY-1998) to the EMBL/GenBank/DBD databases.
DR EMBL: X96913; CAA5632.1; -
DR EMBL: AC004728; -; NOT_ANNOTATED_CDS.
DR FlyBase: FBgn0016122; Acer.
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR00130; Zn_MTPeptide.
DR Pfam: PF01401; Peptidase_M2; 1.
DR PRINTS: PR00791; PEPTIDPTSEA.
DR PRODOM: PD004184; Peptidase_M2; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
FT CONFLICT 236 L -> F (IN REF. 2).
FT CONFLICT 341 Q -> H (IN REF. 2).
FT CONFLICT 528 V -> A (IN REF. 2).
SQ SEQUENCE 630 AA; 73042 MW; 2484AB93139C32D CRC64;

Query Match 24.08; Score 1030; DB 5; Length 630;
Best Local Similarity 35.68; Pred. No. 8.3e-67;
Matches 219; Conservative 113; Mismatches 260; Indels 24; Gaps 10;

QY 6 WLLSLVAVTAASSTIEQAKTFLDKFNHEADLFYQSSLASNYNTITEENVQMNNA 65
DB 16 WLPRLGISMNGSCASYLE-ARRFELENEDELRRRPFHEPLSGNNTVNTTEARQMTIEV 74
QY 66 GDMKSAFLKQSTLAQNPLOETQNTLVKLQALQONSSVLSDEKSKRLNTILTMST 125
DB 75 YARNNAELKRLAQIKSSDYQSEADIRQAEHLKSLGASALNADYIALQNAISSMOT 134
QY 126 IYSTGWCPNDPOEC-LLEPELNLIMANSLDYNNRLMAWSEKSEVKQRLPLEEYV 184
DB 135 NYATAVCTSTNNSDCLTEPHIQERLSRDPALAWYWRWMDKSTPMKONFAEYV 194
QY 185 VLKEMARNVHEDYDYGVDYGVYDYSRGQLLEDVHEPFEETKPYENHNAV 244
DB 195 RLRRKASQNLGHSYADYVQFE-----DPPER-----QDATFTKQLPLRYRLHGY 244
QY 245 RAFLKMANV-PSYISPIGCLPAHLGDMGRFMTNYSLTVPFGQKPNIDVTAMVDAWD 303
DB 245 RFLRQHYGPDVWAPAGNIPISILGMMGQSNWELDLTPPEKRFVYVAKMEKQY 304
QY 304 AORIFPEAKRFYVGLPMWTOGFENSLTDPGNAVQKAVCHPTAMDGLK-GDPRILMCT 362
DB 305 VQLEFLGDFQFQSLGMRALPSPFNNLSVLTLPDD-ROVVCASAMDFQDSVIRIKMT 363
QY 363 KYTMDFLTAHHEMGHIQYDMAUYAOPFLLRNGANGFEEAGELMISTLSAARPKLKSG 422
DB 364 EVDSHFYVYVHHELGHIQYLOYEDQPAVYKRAPRNGHEAVGDIYALSWSAKTLKAG 423
QY 423 LLSPDFQENETELNFKLQALTYGLTPTMYLEKRWVFKGEIPKDDMKKWKMEKR 482
DB 424 LIE-NGLRDEKSRINQLEQKSLIVLEPFGYAVDXYKRAVFRNLDSQNMCGWQARS 482
QY 483 EIVGVVEPRHDETCODASLEFHVNDSEFTYTYRTLYKOFQFQALQAAKHGEP----- 538
DB 483 EFGGVPEPRFTEKDFDPKAKYHIDADVEYLYEFAAHIFQFQFHVLCRKKAQYAPNNSR 542
QY 539 --LHHCDSINSTEAGOKLFNNMLRLGKSEPTWLTLENVYGANMNNRPLANTFEPLTYLK 596
DB 543 LTLDMODLFGSKAAGRSLSQPLSKNSNMKWEVLEEFTGETEMDPAALLEEYEPLOYWK 602
QY 597 DONKNSFVGMSTWSP 612
DB 603 QE--NSRLGVPLGMP 616

```

```

RESULT 15
ID 09VLJ6 PRELIMINARY; PRT; 630 AA.
AC 09VLJ6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ACER protein (U028328P).
GN ACER OR CGI0593.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731133;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazet R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Gelbart W.M., Glasser K.,
RA Fostier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kienison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinett K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sriden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gbbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBD databases.
DR EMBL: AE003621; AAF52693.1; -
DR EMBL: AY051750; AAK93174.1; -
DR MEROPS: M02.002; -
DR FlyBase: FBgn0016122; Acer.
DR InterPro: IPR001548; Peptidase_M2.
DR InterPro: IPR00130; Zn_MTPeptide.
DR Pfam: PF01401; Peptidase_M2; 1.

```



DR PRINTS: PR00791; PEPTIDPTASEA.  
 DR Prodom: PD004184; Peptidase\_M2; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; UNKNOWN 1.  
 SQ SEQUENCE 630 AA; 73057 MW; 6D9355EB5773289 CRC64;

Query Match 24.0%; Score 1028; DB 5; Length 630;  
 Best Local Similarity 35.6%; Pred. No. 1.2e-66;  
 Matches 219; Conservative 112; Mismatches 261; Indels 24; Gaps 10;

QY 6 WLISLVAVTAAGSTIEGQATFLDKFNHEDELFTYSSLASWYNTNTEENVQNNNA 65  
 DB 16 WLPGLSNGNSCASVLE-ARFPELNEQRLRRFHEEPFSGVYNNVTNTEANQAMIEV 74  
 QY 66 GDKNSAFLEQSTLAOMYPLQEIONTLVKLOLALQONSSVLSSEDKRLNTLNTMST 125  
 DB 75 YARNAELNKRALAQOIKSSDYQSDADIRQAEHLKSLGASALNADYIALQNNASSMOT 134  
 QY 126 IYSTGKVCNPDNPEC-LLEPGLNETMANSLDYNERLWAMESWSEVQKOLRPLYEEV 184  
 DB 135 NYATATVCSYTNRSDCSLTLEPHIQERLHSRDPALAWYWRHHDKSGTPMQNFAEYV 194  
 QY 185 VLKEMARAHNYEDYGYWRGDYVNVGDYDSRGQIEDVETFEETPLVEHLHAYV 244  
 DB 195 RLTKRASQLNGHRSYADYWOYE---DDEFER---QLDATERKQLPFYKQHLGYV 244  
 QY 245 RAKIMNAY-PSYISPIGCLPAHLLDGMGRFNTLSLTVPEGOKPIDYTDAMVDQAMD 303  
 DB 245 RFRLRQHYGRPYMAEGNIPISLIGNMWGSNELLDFPYPKRPVDYKAEMEKOYGT 304  
 QY 304 AQRIFKAERKEFVSGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDIGK-GDFRILMCT 362  
 DB 305 VQKLFELGDOFFOSLGRALPPSEFMNLSVLRPDD-RHVVCHASAMFYDDSDVRIKMT 363  
 QY 363 KVTMDDELTAHHEGHIOYMAAOPFLRNGANEGFHRVAGFIMLSATKHLKSTIG 422  
 DB 364 EVDSHYTVVHHEGHIOYLYQEQPARYRGADNPGEHVAAGDYVALSWASKHLKAIG 423  
 QY 423 LLSPEQEDNETELNFKALITVGLPFTYMLEKRWMMVFGEIPIKDDMMKMMEMKR 482  
 DB 424 LIE-NGRLDEKSRINQFLKQALSKTYVLPFGYAVDKRYAVFRRELDSEONGCFWOMRS 482  
 QY 483 EIVGVPEVPHDETYCDPASLHVNSDYSFIRYTRTLVYQFOEALQQAHEGP--- 538  
 DB 483 EFGGVPEVPHDETYCDPASLHVNSDYSFIRYTRTLVYQFOEALQQAHEGP--- 538  
 QY 539 --LHKCDINSTAGCKLFNMLRKGSEPTLAEENVGAKNNVRLVYFPEPLTWLK 536  
 DB 543 LTLNCDIFGSKAAGRLSLOFLSKGNSRHKVEYLEEFTGETENDPALLEYFPEPLYOMLK 602  
 QY 597 DQNKSEFGWSTWSP 612  
 DB 603 QE--NSRLGVPLGWGP 616  
 RESULT 16  
 Q9D836 PRELIMINARY; PRT; 249 AA.  
 ID Q9D836  
 AC Q9D836;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DE 2010305L05RIK protein.  
 GN 2010305L05RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C5/BL/6J; TISSUE=SMALL INTESTINE;  
 RX MEDLINE=21085660; Pubmed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuenl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L. M., Stablil R., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,  
 RA Brownstein M. T., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kamuya M., Lee N. H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K. P.,  
 RA Suzuki H., Toyokawa K., Wang K. H., Weltz C., Whitaker C., Wilmink L.,  
 RA Wyszynski B. A., Yoshida K., Hasegawa Y., Kawai H., Kohlschki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK008530; BAB25723.1; -;  
 DR MEROBS: M02.006; -;  
 DR MGI:1917258; 2010305L05RIK.  
 DR InterPro: IP001548; Peptidase\_M2.  
 DR Pfam: PF01401; Peptidase\_M2; 1.  
 SQ SEQUENCE 249 AA; 28379 MW; 19372B2B78AAE921 CRC64;

Query Match 23.4%; Score 1004; DB 11; Length 249;  
 Best Local Similarity 75.1%; Pred. No. 1.7e-65;  
 Matches 187; Conservative 31; Mismatches 31; Indels 0; Gaps 0;

QY 557 MRLGKSEPTLAEENVGAKNNVRLVYFPEPLTWLKDNKNSFGWSTWSPADQ 616  
 DB 1 MSLGSEPTLAEENVGAKNNVRLVYFPEPLTWLKDNKNSFGWSTWSPADQ 60  
 QY 617 SIKVRLSLKALGDKAYENNDDMYLFSSVAYAMROFLKAYKQMLFGEEDVAVANIK 676  
 DB 61 SIKVRLSLKALGDKAYENNDDMYLFSSVAYAMROFLKAYKQMLFGEEDVAVANIK 676  
 QY 677 PRISFEPTAPKNVDIIIPREVEKATMRSSRIINDARLNDNSTEPLGIPTGPNNQ 736  
 DB 121 PRVSEFFPTSPONSDVLRSEVEDALVMSGRINDVGLNDNSTEPLGIPTGPNNQ 180  
 QY 737 PVSIMLIVGVGVGVIVGIVILFTGIRDRKKRKAKSGENPYASIDISGENNPFQ 796  
 DB 181 PPTIMLIIIFGVMLLVYGIILITVGTIGKRRKKNETRENSVDSMDIGGSENAFQ 240  
 QY 797 NDDVOFTSP 805  
 DB 241 NSDDAQTSP 249  
 RESULT 17  
 Q8SXX2 PRELIMINARY; PRT; 844 AA.  
 ID Q8SXX2  
 AC Q8SXX2;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DE 2010305L05RIK protein.  
 GN RH06639P.  
 GN ANCE-3.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Mirande A., Mungall C. J., Nunoo J., Pacleb J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S. E., Rubin G. M.,  
 Celniker S.;



RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF075532; AAL68339.1; -  
 SQ SEQUENCE 844 AA; 98243 MW; 46E1E123F5CC9F71 CRC64;  
 Query Match 21.7%; Score 930; DB 5; Length 844;  
 Best Local Similarity 33.2%; Pred. No. 2,6e-59;  
 Matches 194; Conservative 119; Mismatches 248; Indels 24; Gaps 9;

QY 46 ASNNYNTNTEENVQNMNNAQKWSAFLKEQSTLQMTYLOLQNTVLTOLQALQONS 105  
 DB 234 AOMNFETVNDFTQOTALNAQRYVEFORITAEQSKRIKDLFDRRLRYOLMOSEVGP 293  
 QY 106 SVLSEDSKRLNTILNTMSTIYTGKVCNPDNPOEC-LILEPGLNEIMANSIDYMERLMA 164  
 DB 294 NALPLDVRNRYNLLNEMFLVYNSAETICAYQOPQCDLHYITQLKDIAMKSDMDELQHT 353  
 QY 165 WESWSEVQKOLRPLYEEVYVLKNEBARNHVEDYDGRGDEYVNGVDGYTSRQOLTE 224  
 DB 354 WVEYHRRAGSGMRDSEQLIDWQEVAVYNNVTNGEYWTAYE-----SGNFRQ 403  
 QY 225 DVEHTEETKPLTEHHAAYVRAKIMAY-PSYISPIGCPAHLIDGMGRFNTLXSLTV 283  
 DB 404 DMQIYEQILPTEGLHAAYVRKLDYGPDRINRIAPISHLGNMTGQSMNVLDLIT 463  
 QY 284 PEGQKNIDVTDAVDQADQRIPEKAEKFEVSVGLPMTGQFWMENMLTPGNVQKAV 343  
 DB 464 PYGRKLIDVTPRVWQGTPLQMPQADPEFFTSINMSAVGEFRNSLFEQPLD-RVYL 522  
 QY 344 CHTPLANDL-GKDFRLMCTKYTMDPLTAHHEMGHIQDYAMAAPPELLRNGANGEGHE 402  
 DB 523 CEESASDFCNHRDFRVKICTDINORSLSIVHHEMAHIQVFLQYRHLPIKIFRNGANPAFRQ 582  
 QY 403 AVEISLSAAPPKHLKLSLSPDQENMEINFLQALITVSTPLFTMLEKMRM 462  
 DB 583 ANQDAIGLSVSTPRHLQTLGLORSLDE-SSDIDINFLMALDKAFLDPAFLSLDWRD 641  
 QY 463 VFKEGELPKDQMKMKMKREIYGVVEPYPHDETYCDPASLPHVSNDSYFIYRTTLYQ 522  
 DB 642 VEGSNANKRPMNCHYNNLEKYSIGIKRPLRSEKDFDPGAKHIANINIIYIFPSTVLD 701  
 QY 523 FQEQELCOQAKH-----EGPLKCDISNSTEAGOKLENNMLRLGKSEPTWLTALENVGA 576  
 DB 702 FOYLRGLCRSGQYVGPDRKPLHQCIDIROPAGNLTILKLSKSGASQPMQVLEETLRE 761  
 QY 577 KNNVAPRLNLYPEPLFTMLKDON-KNSFVGSTWMSPYADSIK 619  
 DB 762 GRLDGTALREYFALEEMLRQENLRITNEYGNMYP-GDYCKRSIE 805

RESULT 18  
 09VJVI PRELIMINARY; PRT; 792 AA.

AC 09VJVI  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE CG17988 protein.  
 GN ANCE-3 OR CG17988  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Eukaryota; Metazoa; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidae; Drosophilidae; Drosophila.  
 OX NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brannon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Bens P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,  
 RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Centier A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,  
 RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 RA Glodex A., Gong F., Gotrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,  
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Minkulov G., Mishina N.V., Mobarly C., Morris J., Moshirei A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson W., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AF003641; AAF53356.1; -  
 DR FlyBase; FBgn0032536; Ance-3.  
 DR InterPro; IPR001548; Peptidase.M2.  
 DR Pfam; PF01401; Peptidase.M2; 1.  
 DR PRINTS; PR00791; Peptidase.M2.  
 DR PRODOM; PD004184; Peptidase.M2; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN.1.  
 SO SEQUENCE 792 AA; 92321 MW; 757EC472089022B0 CRC64;

Query Match 19.6%; Score 841; DB 5; Length 792;  
 Best Local Similarity 30.5%; Pred. No. 7,8e-53;  
 Matches 196; Conservative 109; Mismatches 224; Indels 118; Gaps 14;

QY 48 WNTNTNTEENVQNMNNAQKWSAFLKEQS-TLAQMT-----PLQETQNTYK 94  
 DB 158 YNNNNPNVVEFV-GINNRRFENPFLSNODRFNLQGLERORYQDRRYQDELEKRLIL 216  
 QY 95 L-----OLQALQONGSSV 107  
 DB 217 LVESDQKSELECTANLNAQRYVEFORITAEQSKRIKDLFDRRLRYOLMOSEVGPNA 276  
 QY 108 LSEDSKRLNTILNTMSTIYTGKVCNPDNPOEC-LILEPGLNEIMANSIDYMERLMAWE 166  
 DB 277 LPLDVRNRYNLLNEMFLVYNSAETICAYQOPQCDLHYITQLKDIAMKSDMDELQHTW 336  
 QY 167 WRSSEVQKOLRPLYEEVYVLKNEBARNHVEDYDGRGDEYVNGVDGYTSRQOLTEY 226  
 DB 337 EYHRRAGSGMRDSEQLIDWQEVAVYNNVTNGEYWTAYE-----SGNFRQDM 386  
 QY 227 EHTFEETKPLTEHHAAYVRAKIMAY-PSYISPIGCPAHLIDGMGRFNTLXSLTVP 285  
 DB 387 DIWEOILPTEGLHAAYVRKLDYGPDRINRIAPISHLGNMTGQSMNVLDLITP 446  
 QY 286 GQKNIDVTDAVDQADQRIPEKAEKFEVSVGLPMTGQFWMENMLTPGNVQKAVCH 345  
 DB 447 PGRKLIDVTPRVWQGTPLQMPQADPEFFTSINMSAVGEFRNSLFEQPLD-RVYLCE 505  
 QY 346 PTANDL-GKDFRLMCTKYTMDPLTAHHEMGHIQDYAMAAPPELLRNGANGEGHEAV 404  
 DB 506 PSAMDFCNHRDFRVKICTDINORSLSIVHHEMAHIQVFLQYRHLPIKIFRNGANPAFRQAV 565

```

OY 405 GEIMSLATPKHLKSLGSLSPDFQEDNETETINFLKQALITVGLPTVMLEKRMNV 464
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 566 GDAIGLSVSTRHLOTLELORSIDE--SYDINTLFTMAIDKAPALSLDMWRDVF 624
OY 465 KGEIPKDKMKRMKREIYGVVEPVHDETYCDPASLFHVSNDYSFIRYTYTLQFO 524
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 625 SGNARKRTMNCWYN-----LRFESTVYQFO 651
OY 525 FOEALCOAAKH-----EGPIHKCDISNTEAGOKLEMLRGSEPTTALENVGAKN 578
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 652 IYRGCSRSGGYVPGDPKPLHCDIYRQPAAGNMLTKMSKASOPWQOEVLTEBGR 711
OY 579 MNVBPLLYEPLFTWLDKQDN--KNSFYGWSTDMSPYADOSIK 619
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 712 LDGTALREYFAPLEBMLRQENLRINEYGVWYD-GDYCKRSIE 753

RESULT 19
Q18581 PRELIMINARY; PRT; 907 AA.
AC 018581:
DT 01-NOV-1996 (Tremblrel, 01, Created)
DT 01-NOV-1996 (Tremblrel, 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE Hypothetical 101.1 kDa protein.
GN C42D8.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX Hallsworth K.;
RT "The sequence of C. elegans cosmid C42D8.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: U56966; AAA98719.1; -.
DR InterPro: IPR001548; Peptidase_M2.
DR Pfam: PF01401; Peptidase_M2; 1.
DR ProDom: PD004184; Peptidase_M2; 1.
KW Hypothetical protein.
SQ SEQUENCE 907 AA; 101086 MW; F9B8B5C0F9BC5AA3 CRC64;

Query Match 15.0%; Score 642.5; DB 5; Length 907;
Best Local Similarity 27.0%; Pred. No. 3.1e-38;
Matches 172; Conservative 136; Mismatches 289; Indels 41; Gaps 19;
OY 2 SSSSMLLSLVAVTAOSTIEQAKTFLDFNHEADLFYSSSLASWNTTTEENVGN 61
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 SSNWTKNDNLOAPGSIKD--EKIKRSLAGYEAEIKVLRVALSGRIYNDKSPSLKLA 217
OY 62 MNNAGDKWSAFLEKOSTLAOMYLOETIQLNTVKIQLQALOQNGSSSVLSEDKSRRLTILN 121
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 218 LDEANVNLTFVRSSTQAKOFDMASVTDEKVRKOLGYVSFEQMSALAPSRFADYSQAQA 277
OY 122 TMSIYTGKVCNDNQECILLEPGNEMANSLOYNEMLMWESKRSRSGVQLRPLYE 181
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 278 ALNDSKDSITCDKDVPPPCALQKIDMDSIFRNEKDSASRLQHLWVSIVTAIAKS-PPSYN 336

```

```

OY 182 EYVVLKEMARAHNYEDYCTYMGDYEVNG-VGCTYYSRQGLIEDVEHFEETKPLYEH 240
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 337 NITTSNEGAKLNGFANGGAMRSAPDMSKVHAEF---DLNKQIDKITYSTIQPYQL 393
OY 241 HAVVRKAKIMAY--PYSISPGLPAHLIGDMGGRWTNLSLTVEFGGKPNIDYDAV 298
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 394 HAYMRQLAGIYSNPVGLSKDGPPIPAHLFGSLDGGWMSAHYEQTKEEES--ETPEAML 451
OY 299 D----QAMPQRITFEKAEKFEVSVGLPNTQGEFWSMLTDGNYQKAVCH--TAMD-LG 352
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 452 SAVNTQWYTTKKMEVATYAFYFKSAGPPLPKSVWTSIARWVS-KDMCHPAALADMA 510
OY 353 KGPRLIMCTKYVMDPFLTAHHEMGHIQYDMAAOPFLIRNGANGDFHEAVETKLSA 412
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 511 PNDFRVACAOQLGEPPFEQASHLVOTYQYLYTKDQSLFRQASPVITDAANFAHLS 570
OY 413 ATPKHLKSLGSLSPDFQEDNETE-INFLKQALITVGLPTVMLEKRMNVFKGIPKD 471
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 571 TNPHYLSOKLYPSEHLIDKDSYIINKLYKESLESTFKPLTIAADNWKYELFDGTVPKN 630
OY 472 QMKKMKMKREIYGVVEPVHDETYCDPASLFH--VSNDYSFIRYTYTL-----YQFOF 525
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 631 KLNDRMWELRNKYEGRSPQYNTSMID--ALIHNSVQVHS---PATRTLSYLVKFOI 685
OY 526 QEALCOAA--KHGPIHKCDISNTEAGOKLEMLRLKSEPTTALENVGAKNMVNR 582
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 686 LKALCORELFWSESG---CLISEDTT--EKRETKLSSITWKLALMISGKGLDAQ 739
OY 583 PLNVYEPFLFTWLDKQDN--NSFYGWSTDMSPYADOSI 618
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 740 PLEYEPLNLRNNEIDQYVVGWDSGETFYVEI 777

```

```

RESULT 20
Q9V520 PRELIMINARY; PRT; 661 AA.
AC 09V520:
DT 01-MAY-2000 (Tremblrel, 13, Created)
DT 01-MAY-2000 (Tremblrel, 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel, 21, Last annotation update)
DE CG8196 protein.
GN ANCE-4 OR CG8196.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY.
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferreira C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```





```

Db      61  NMNNAADKWSAFLEQSTIAOMPILOEIONLTVKLOLQALQONGSSVLSDEKSKRLNTLT 120
QY      121  NTNSTYTGKVCNPNPOECILLEGLMEIANSIDYERLWAMESRSSEVGKOLRPLY 180
Db      121  NTNSTYTGKVCNPNPOECILLEGLMEIANSIDYERLWAMESRSSEVGKOLRPLY 180
QY      181  EEEVVLKNEMARANHYEDYDGYWAGDYVSDYSGQLIEDVEHFEETIKPLYEHL 240
Db      181  EEEVVLKNEMARANHYEDYDGYWAGDYVSDYSGQLIEDVEHFEETIKPLYEHL 240
QY      241  HAYVRKLMNAYPSYISPGCLPAHLGDMGREGMTNYSLTVPFGQKNIDVTAMVDO 300
Db      241  HAYVRKLMNAYPSYISPGCLPAHLGDMGREGMTNYSLTVPFGQKNIDVTAMVDO 300
QY      301  AMDAORIFKEAEKFEFVSGLPMTQGFENSMULTDPGNOKAVCHPTAMDLSKGFRLM 360
Db      301  AMDAORIFKEAEKFEFVSGLPMTQGFENSMULTDPGNOKAVCHPTAMDLSKGFRLM 360
QY      361  CTKVMTMDFLTAHHEMGHIOYDMAVAAPFLLRNGANGFHEAVGEIMSLAATPKHLKS 420
Db      361  CTKVMTMDFLTAHHEMGHIOYDMAVAAPFLLRNGANGFHEAVGEIMSLAATPKHLKS 420
QY      421  IGLLSPDQEDNETINFLKQALITVGTLPPTYMLEKRWMTKEGELPKDQMKKWMEM 480
Db      421  IGLLSPDQEDNETINFLKQALITVGTLPPTYMLEKRWMTKEGELPKDQMKKWMEM 480
QY      481  KREIVGVEPVPHDETYCDPASLFHVSNDSYFIRYRTLYOFQFOEALCOAAHHEGFLH 540
Db      481  KREIVGVEPVPHDETYCDPASLFHVSNDSYFIRYRTLYOFQFOEALCOAAHHEGFLH 540
QY      541  KCDISNSTEAGOKLFNMLRLGKSEFWTLAENNVGAKNMNVRPLNTEFELTWLKDONK 600
Db      541  KCDISNSTEAGOKLFNMLRLGKSEFWTLAENNVGAKNMNVRPLNTEFELTWLKDONK 600
QY      601  NSFVGMSTWSPYADQSKVIRISLSKALGDKAYEMNEMNEMLEFSSVAYAMROYFLVKN 660
Db      601  NSFVGMSTWSPYADQSKVIRISLSKALGDKAYEMNEMNEMLEFSSVAYAMROYFLVKN 660
QY      661  QMILFGEEDVAVANLKPRISFNFVTAPKNVSDIIPTEVEKATIRMSRSTRINDAFRLND 720
Db      661  QMILFGEEDVAVANLKPRISFNFVTAPKNVSDIIPTEVEKATIRMSRSTRINDAFRLND 720
QY      721  SLEFLGIOPTLGPNOPPVSTWMLVGEVGVYVIGVILITGIRDKKKNNARSGENP 780
Db      721  SLEFLGIOPTLGPNOPPVSTWMLVGEVGVYVIGVILITGIRDKKKNNARSGENP 780
QY      781  YASIDISKENNPGFQNTDDVQTSF 805
Db      781  YASIDISKENNPGFQNTDDVQTSF 805

```

```

RESULT 2
US-08-481-626-2
; Sequence 2, Application US/08481626
; Patent No. 5801040
; GENERAL INFORMATION:
; APPLICANT: Soudrier, Florent
; APPLICANT: Alenc-Gelas, Francois
; APPLICANT: Hubert, Christine
; APPLICANT: Corvol, Pierre
; TITLE OF INVENTION: Nucleic Acid Coding for the Human
; TITLE OF INVENTION: Testicular Angiotensin Converting Enzyme (ACE) and its
; TITLE OF INVENTION: Uses, Especially for the In Vitro Screening for this
; TITLE OF INVENTION: Enzyme in the Organism
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,626
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/656,183
FILING DATE: 04-MAR-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 89-09062
FILING DATE: 05-JUL-1989
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 04958-0006-02000
TELEPHONE: 202-408-4400
TELEFAX: 202-408-4400
INFORMATION FOR SEO ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 732 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-481-626-2

```

Query Match 31.3%; Score 1344; DB 1; Length 732;

Best Local Similarity 41.8%; Pred. No. 4,8e-120;

Matches 259; Conservative 119; Mismatches 204; Indels 38; Gaps 10;

```

QY      15  TAAOS-----TIEQATFLDKFHEAEDLFYOSLSAMNNTNTEE-----NQNM 62
Db      61  TSAOSPNLTDEAKSKFVEEDRTSQVYWNELAEAMNNTNTEESKILLQNMIOA 120
QY      63  NNAAGDKWSAFLEQSTIAOMPILOEIONLTVKLOLQALQONGSSVLSDEKSKRLNTLT 122
Db      121  NHT-----LKYGTQAKKFDVNOLONTTIKRIIKYQODDERALLPPOLEETKILLD 172
QY      123  MSTYTGKVCNPNPOECILLEGLMEIANSIDYERLWAMESRSSEVGKOLRPLYEE 182
Db      173  METTYSVATVCHNG--SCLQLEPDLTNVMASTRKYEDLLWAMEGRDKAGRAILQFYPK 230
QY      183  YVVLKNEMARANHYEDYDGYWAGDYEVNGVDGYDSRGQLEDVHEHTEETIKPLYEHL 242
Db      231  YVELINQARLNGYDADGWSMSETEPSLE-----QDLERLQELQPLYLNLHA 280
QY      243  YVRKLMNAY-PSYISPGCLPAHLGDMGREGMTNYSLTVPFGQKNIDVTAMVDO 301
Db      281  YVRALHRRHYGQHINLEBPIPAAHLGMMMAQTWENIYDLVPPSPASMDTTEMLKQG 340
QY      302  WDAORIFKEAEKFEFVSGLPMTQGFENSMULTDPGNOKAVCHPTAMDLSKGFRLM 360
Db      341  WPRRAMEKADDEFTSLGLLPVPPEFWKSMLEKPTDREEVNCHASADFYNGDFRIKQ 400
QY      361  CTKVMTMDFLTAHHEMGHIOYDMAVAAPFLLRNGANGFHEAVGEIMSLAATPKHLKS 420
Db      401  CTTVMLDELVAHHEMGHIOYDMOKDPAVLREGANGFHEALGDVLAISVTPKHLHS 460
QY      421  IGLLSPDQEDNETINFLKQALITVGTLPPTYMLEKRWMTKEGELPKDQMKKWMEM 480
Db      461  IMLLSSEGSD-EHDINFLMKMALDKIAFIPFSYLVQMRWRVDSITKENYNOEWMSL 519
QY      481  KREIVGVEPVPHDETYCDPASLFHVSNDSYFIRYRTLYOFQFOEALCOAAHHEGFLH 540
Db      520  RLKYGGLCPVPYPRQGDPPCAKTHISSVYIYFVSIIQFOFHALCOAGHGTPRLH 579
QY      541  KCDISNSTEAGOKLFNMLRLGKSEFWTLAENNVGAKNMNVRPLNTEFELTWLKDONK 600

```

Db 580 KCDIYQSKGQRLATAMKIGFSRPMPEAMQILITGPPNMSASAMLSYFPLDMLTENE 639  
 QY 601 --NSFVGM-STDWSPYADOS 617  
 Db 640 LHGEKLGWPOYNWTPNSARS 659

## RESULT 3

US-08-989-299-4  
 ; Sequence 4, Application US/08989299  
 ; Patent No. 6194556  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Acton, Susan L.  
 ; APPLICANT: Robinson, Keith E.  
 ; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG  
 ; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
 ; STREET: One Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109-2170  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/989,299  
 ; FILING DATE: 11-DEC-1997  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Arnold E., Beth  
 ; REGISTRATION NUMBER: 35,430  
 ; REFERENCE/DOCKET NUMBER: MIA-025.01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-832-1000  
 ; TELEFAX: 617-832-7000  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 732 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-989-299-4

Query Match 31.3%; Score 1344; DB 4; Length 732;  
 Best Local Similarity 41.8%; Pred. No. 4, 8e-120;  
 Matches 259; Conservative 119; Mismatches 204; Indels 38; Gaps 10;

QY 15 TAAOS-----TTEBOAKTFLLDKFNHEADLFYQSSLASWYNNTNITE-----NVQNM 62  
 Db 61 TSAOSPMLYVDEAKSFEVEYDRTQSVNNETAEANMNNNTNITETSKILLQNMQA 120  
 QY 63 NNADGKMSAFLEQSTLAQMYPLQELQNLTKLQALQONGSSVLSDEKSKRLNTLNT 122  
 Db 121 NHT-----LKYGTQARKFEDVNOQLNTIKRIKKVQDLERAPALPAQELREYNKILLD 172  
 QY 123 MSTYSGTKVCNPDPNPOECILLEGELMEIANSIDYNERLMAWESRSEVQKQRLRPLYEE 182  
 Db 173 METYSVATVCHPNQ--SCDLEPDLNVMATSKYEDLLMAWEGWDRKAGRAILOFYRK 230  
 QY 163 YVVLKNEARAHNEDYDGRGDEYVNGVGYDYSRGQLEIEVEHTFEERIKPLYEHLA 242  
 Db 231 YVELINQAKRLNGYDADGSRMSYETPISLE-----QDLERLFOELQPLYLNLIR 280  
 QY 243 YVRAKLMNAY--PSYISPGICLPAHLGLDMGKREFTNTLSLTVPFGQPNIDIVDAMDQA 301  
 Db 281 YVRALHRYGAQHINIEGPPIAHLILGNMAQTSNITDIYLVPPSAPSMDTTEAMLKQG 340

QY 302 WDAQRIKKEAEKFFVSYGLPNMTQGWENSMLEDPGNVOKAVCHPTAMDLGK--DEFILM 360  
 Db 341 WTPRRMFKEADDEFTTSIGLLPPEPEFWNKSMELEKPTDGEVYVCHASANDFYGKQFRKQ 400  
 QY 361 CTKVTMDPFLAHHEMGHIOYDMAAOPFLLRNGANEGEHEAVGEIMSLNATEPKHLKS 420  
 Db 401 CTYNLEEDLYVAHHEMGHIOYFMQYKDLPVALLREGANPGFHEAIGDVALYSTPKHLHS 460  
 QY 421 IGLSPDQEDNETEINFLKQALTYGLPFTYMLEKRWYKKEIPEKQDMKKWEM 480  
 Db 461 LNLSSGSGSD--EHDINFLMKMALDKIAFIPFSYLDQWKRVRFGDSITKENINDENMSL 519  
 QY 481 KREIVGVEPYPHDETCDPASLEFVENSQYSTRYTRTLTYQFOFQALQQAQKHEGPLH 540  
 Db 520 RLKYQGLCPVPRTGDPDFQAKFIHSSVPYIRYVPSFIQGFHEALQQAAGHTGPLH 579  
 QY 541 KCDISNSTEAGOKLFNNMLRKGSEPTWLALENVGAKNMVRPLNFEPLFTWLDQMK 600  
 Db 580 KCDIYQSKGQRLATAMKIGFSRPMPEAMQILITGPPNMSASAMLSYFPLDMLTENE 639  
 QY 601 --NSFVGM-STDWSPYADOS 617  
 Db 640 LHGEKLGWPOYNWTPNSARS 659

## RESULT 4

US-08-989-299-7  
 ; Sequence 7, Application US/08989299  
 ; Patent No. 6194556  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Acton, Susan L.  
 ; APPLICANT: Robinson, Keith E.  
 ; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG  
 ; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
 ; STREET: One Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109-2170  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/989,299  
 ; FILING DATE: 11-DEC-1997  
 ; CLASSIFICATION: 514  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Arnold E., Beth  
 ; REGISTRATION NUMBER: 35,430  
 ; REFERENCE/DOCKET NUMBER: MIA-025.01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-832-1000  
 ; TELEFAX: 617-832-7000  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1306 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-989-299-7

Query Match 31.2%; Score 1337; DB 4; Length 1306;  
 Best Local Similarity 41.7%; Pred. No. 5, 9e-119;  
 Matches 255; Conservative 118; Mismatches 204; Indels 34; Gaps 9;

QY 20 TTEBOAKTFLLDKFNHEADLFYQSSLASWYNNTNITE-----NVQNMNAGDKWKA 71

```

Db 644 TDEAFASKFVEEYDRTSQVWVMEYAEAMNNTNTTTSTKILLQKNQJANH----- 697
QY 72 FLKEOSTLAQMPLOEIQNLVYKLOLQOQSSVLSDESKRLNTLNTMSTYSTGK 131
Db 698 --LKTGTQARKKDVNOLOWTIKRIKKVQDLERFALPAQLEENKILLDMETTYSAT 755
QY 132 VCNPNPQPCLLLEFGLMEINANSIDYERLMAESMSSEVQKQRLPYEYVVLKEMA 191
Db 756 VCHPNG--SCQLERPDILNVATSKRYKIDLLAMEGMDKGRALILQFYPRYVELIQQA 813
QY 192 RANHVEDYDGYRKYEVNGVDYDSRGQLEDEHTFEERIKPLEYHLAAYRAKLMA 251
Db 814 RLVGYDAGDSWRSMYTERPSLE-----QDLERLFOQLYLNLAAYRRALHRH 863
QY 252 Y-PSYITGCLPAHLGLDMGREFWNLVSLTFPQGRPNIDVTDAVDAQDAQRIEKE 310
Db 864 YCAQIINLEGP1PAHLGONMAQWTSNITDLVPEPSAPMDTTEALMKQGMTPRIMFE 923
QY 311 AEKPFVSGLEPMTQGEWENSLTDPGNVQKAVCHPTAMDLGK-DPRILMCTKYTMDP 369
Db 924 ADDEFTSLGLLPVPEPMNKSMLKPTDGREVVCHASAMDFYNGKDFRIKQCTVLEDL 983
QY 370 LTAHEMGHIQDMAAQAAPFLRNGANEGFHEAVGEIMLSAATKHLKSLGSLSPDQ 429
Db 984 VVAHEMGHIQYFMQYKDLVVALREGANPFGHEAIGVLAHSVTPKHLHSLNLSSEGG 1043
QY 430 EDNETEINFLKQALITVGLPPTVYMLEKRWYFEGELPKQOMKMKWEMKREIYGYE 489
Db 1044 SD-SHDINFLMKALDLIAIPSYLVQWRMRFVDSITKENYNGEWSMLXKQGLCP 1102
QY 490 PVPHDEYCDPASLFHYSNDYSFIRYTRTLVQFOFQALCOAKHEGLKCDISNSTE 549
Db 1103 PVPRQGDQDFPGAKFHIPSSVPYIRIFVFTIOFQHEALCOAAGTGLHCKDIYOSKE 1162
QY 550 AGQKFLMLRLKSEWPTLALENVGAKMNVRLINTEPELFTWLKQONK--NSFVGM- 606
Db 1163 AGQRLATAMKLGSRWPEAMQLITQPMNASAMLSYKPLLDMLRTENELHSGKLGMP 1222
QY 607 STWSPYADQS 617
Db 1223 QYMWTPNSARS 1233

RESULT 5
US-08-989-299-5
; Sequence 5, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430

```

```

; REFERENCE/DOCKET NUMBER: MIA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 732 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-989-299-5

Query Match 31.1%; Score 1334; DB 4; Length 732;
Best Local Similarity 42.6%; Pred. No. 4,4e-119;
Matches 255; Conservative 112; Mismatches 213; Indels 18; Gaps 7;

QY 20 TIEQATFLDKFNHEADLFYQSSLASWNYNTNTEENVQNNNAGDKSAFLKQSTL 79
Db 69 TDEAKAARFVEEYDRTSQVWVMEYAEAMNNTNTTTEGSKILLEKSTEVSHTLKYGIR 128
QY 80 AQMTPLQEIQNLVYKLOLQOQSSVLSDESKRLNTLNTMSTYSTGVCVCPDNDPQ 139
Db 129 AKTQDVSNFQNSSIKRIKRLQMDRAVLPKQLEBYNQLILDMETYSLSNICYTN-- 186
QY 140 ECLLEFGLMEINANSIDYERLMAESMSSEVQKQRLPYEYVVLKEMARAHYEDY 199
Db 187 TCMLEPDLNLMMATSKREELLMAMKSWRDKGRALLPFFPKYVEFSKIAKLNGYIDA 246
QY 200 GDYWRGDYEVNGVDYDSRGQLEDEHTFEERIKPLEYHLAAYRAKLMAAYS-YISP 258
Db 247 GDSWRSILYEDNLE-----QDLKLYOELQPLVNLHAAYRRALHRYGSEYINL 296
QY 259 IGLPRAHLGDMGREFWNLVSLTFPQGRPNIDVTDAVDAQDAQRIEKEAEKPFVSY 318
Db 297 DGP1PAHLGONMAQWTSNITDLVPEPSAPMDTTEALMKQGMTPRIMFEADFTSL 356
QY 319 GLPNMTQGEWENSLTDPGNVQKAVCHPTAMDLGK-DPRILMCTKYTMDPFLTAHEMG 377
Db 357 GLLPVPEPMNKSMLKPTDGREVVCHPSAMDFYNGKDFRIKQCTVNMEDLVTAHEMG 416
QY 378 HIQYDMAVAQAAPFLRNGANEGFHEAVGEIMLSAATKHLKSLGSLSPDQENETFEIN 437
Db 417 HIQYFMQYKDLVPTREGANPFGHEAIGDIMALSVTPKHLHSLNLSSTE-GSGYEYDIN 475
QY 438 FLTKQALITVGLPPTVYMLEKRWYFEGELPKQOMKMKWEMKREIYGYEVPYHDETY 497
Db 476 FLTKMALDKIAFIPSSYLDQWRMRFVDSITKENYNGEWSMLKQGLCPVPRSQGD 535
QY 498 CDPASLFHYSNDYSFIRYTRTLVQFOFQALCOAKHEGLKCDISNSTEAGQKLFNA 557
Db 536 FDPGSKFHVPAVRYRVFVFTIOFQHEALCOAAGTGLHCKDIYOSKEGKLLA 595
QY 558 LRLKSEWPTLALENVGAKMNVRLINTEPELFTWLKQONK--NSFVGM-STWSP 612
Db 596 MKLGSKFPEAMQLITQPMNASAMLSYKPLLDMLRTENELHSGKLGMP 653

RESULT 6
US-08-989-299-8
; Sequence 8, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG
; TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA

```



us-09-978-385-2.raii

COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/989, 299  
FILING DATE: 11-DEC-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arnold E., Beth  
REGISTRATION NUMBER: 35,430  
REFERENCE/DOCKET NUMBER: MIA-025.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1312 amino acids  
TYPE: amino acid  
STRANDEDNESS: 1  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
IS-08-989-299-8

[illegible]

-08-989-299-9, Application US/08989299  
Sequence 9, Patent No. 6194556  
GENERAL INFORMATION:  
APPLICANT: Acton, Susan L.  
APPLICANT: Robinson, Keith E.  
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG  
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR  
TITLE OF INVENTION: 14  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/989,299  
FILING DATE: 11-DEC-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arnold E., Beth  
REGISTRATION NUMBER: 35,430  
REFERENCE/DOCKET NUMBER: MIA-025.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-852-7000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1313 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-989-299-9

```

Query Match      30.5%; Score 1310; DB.4; Length 1313;
Best Local Similarity 42.0%; Pred. No. 2,3e-116;
Matches 251; Conservative 115; Mismatches 214; Indels 18; Gaps
7

QY      20  TIEDAQETLDKFNHEHEDLFYQSSLASNNVNTNTTEENVQNMNNAGDKNSAFLEQSTL 79
      1 1 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      650  TDEAANFNEFYEXRYTKAVLMNEYAEANNNHHYNTIESSKILLQKNEVSNHTLKGTV 709
      80  AQMYEIOEIQMLFYKLOLQALQNGSSVEDSKRLNTLMTSTIYSGKVCNPDPQ 139
QY      710  AKTEFVSNFQNSTIKRIKKQVNDRAVLRLPNLEIEYNQILDMETYSVAVCYTNQ-- 767
      140  ECLLLEPGCLNEMASLDYNERLWAMESRSEWGKOLRLPYEYVVLKREMARAHYEDY 199
QY      710  AKTEFVSNFQNSTIKRIKKQVNDRAVLRLPNLEIEYNQILDMETYSVAVCYTNQ-- 767
      140  ECLLLEPGCLNEMASLDYNERLWAMESRSEWGKOLRLPYEYVVLKREMARAHYEDY 199
Db      768  TCIETLEPDLTNMAISKRKYEEELWMMKSWMRKVGRALLPFPFPYVPSNKLAKNGYSDA 827
QY      200  GGYWRGDIYEVNGVDYDYSRGQILDEYHREPEEIKRPLYEHLHAYVRKILMAYPS-YISP 258
      828  GDSWRSYESSDE-----ODLKLQYQELQPLYLNLHAAYRSLRHHNGSEYIML 877
QY      259  IGCPLPAILGDMGRFNTMLSYLVPFGCQPNIDVTYAMNQOADAORIFKEAKEFEVSY 318
      878  DGPILPAHLGNMAAQWMSNIYDLVAFPFSPAPSIDTMTKMTKQCTPRIRIKEDADNFETSL 937
Db      319  GLPNNQTQEWMSNMLDPPGVOKACHPAPMADLGKG-DFRILMCTKYTMDDFLAHHENG 377
      938  GLLPVPEPFVWRSMLKEFYDGRREVYQCHASMDPFYNGKDFRIKQCSYNNMELVLAHHENG 997
QY      378  HIQYDMAVAAQPELLLRGANEGFHEAVGELIMSLAATPRKLSIKLSLSPDQEDNETEIN 437
      998  HIQYMQKDKDPLVTFREGANGFHEAGIDVALSVSTPRKHLSSIMLSSE--GSGYEHDIN 1056
Db

```

Sequence 10, Application US/08989299  
Patent No. 6194556

GENERAL INFORMATION:  
APPLICANT: Acton, Susan L.  
APPLICANT: Robinson, Keith E.  
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG  
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:

STREET: One Post Office Square  
CITY: Boston  
STATE: MA

COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM  
MEDICAL CENTER

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 115 408 4000

FILING DATE: 11-DEC-64  
 CLASSIFICATION: 534

ATTORNEY/AGENT INFORMATION  
NAME: AR0013

NAME: Arnold E., Beth  
REGISTRATION NUMBER: 35,430  
REFERENCE CODE:

REFERENCE/DOCKET NUMBER  
TELECOMMUNICATION INFORMATION

TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000

INFORMATION FOR SEQ ID NO: 10  
SEQUENCE CHARACTERISTICS:

LENGTH: 1310 amino acids  
TYPE: amino acid

STRANDEDNESS:  
TOPOLOGY: 34

MOLECULE TYPE: protein

0T-239-10

query match	29
Best Local Similarity	36

Matches 275; Conservativ

7 LLSLVAVTAAQSTI---

21 LLLLRRPPAATLDPG;

55 TEENVÖNNAGDKWSAFT

81 TAENARQEE-----EAL

107 VLSEDSKRINTT.NTMST

136 NLPLAKROOYNSTI SNWCO

167 SWRSEVCKOI BPI VETUT

[illegible]

1

Db 196 GWHNAVGLPRLPQLQFETALSNEAYRQDGFSPGAYWRSWSDPTE-----EDL  
 QY 227 EHTEBEIKPLLEHLAAVRAKLNNVA--PSYISPICTEPAHLGLGMGREFWTLNLSLTPF  
 Db 246 EHTYQHELEPLNHLAAVRFVLRHRRYGDRTYLRKRPILAHLLGMMMAQSMESIYDVAPE  
 QY 266 GQRPNIDVTAMVDAQMDAORLEPKAEKFEFVSGLPMNTOGFWEBSMLTDPGNVQKACH  
 Db 306 PDKNIDVTSTVQKGMNNTIMEFVABEEFETSLGLTLMPEPEFAESMLEKPDGREVCH  
 QY 346 PTAWDL-GRGDFEILMCTRYTNMDEFLTAHHMHGHIOYDMAVYAAOPELLNNGANBGEFNAV  
 Db 366 ASAMDEYNKRDRFKOCQYQVIMDOLSTVHEHMGHOYLYQKDQPSLRR-ANGFHEAI  
 QY 405 GELMSISAATPKHLISGLTSPDFQEDNETEINPILKOLLTYGTLPTTYMLEKRMNVE  
 Db 425 GDVLTALVSTPRAHKEIGLID-HVTNDESDIYLLKMALEKIALPGLGYLVDDMRGVE  
 QY 465 KGEIPDKOMKMMKMRKEIYGVPEPHDEYCDPASTLFEHNSDYSIRYRTBLTQRO  
 Db 484 SGRTESSRYNDWMYLRKTYQJCGIPYVNHETHFAGAKFHPISVTPYIRYFVSPLQRO  
 QY 525 FOEALCOAKKEGGLHCKDISNSTEAGOKLFMLRLKSEPTLALNNGAVANNMVRPL  
 Db 544 FHOALCMAGHOGPLHOCIDIYSTRAGAKLRAVLGAGCSRPMQVGLDVMVASDALDAPL  
 QY 565 LATEPEPLFWLKDOKNKS--FVGM-STDMSPYADQSIKVRISKSALG-----DKA  
 Db 604 LDYFQPYTWOLOEONBERNEGVLMPEYQMRPLNNPPEGIDLVTDCAEASRFVEEDRS  
 QY 633 YE--WMD-----NEWYLFRRSVAAMQYRLKKNOMILGSEEDVRANLKRISFNEF  
 Db 664 FOAVMNEVYEAANNNTYNTITTEASTILLOKMKQIAHNTLYTG-----MNAERDVSNF  
 QY 685 VTPA-----KNVSD-----ITPREVER 702  
 Db 717 QNATSKRIIRKVDLORAVLPEVKELEE 743  
 RESULT 9  
 US-08-989-299-6  
 Sequence 6, Application US/08989299  
 Patent No. 6194556  
 GENERAL INFORMATION:  
 APPLICANT: Action, Susan L.  
 APPLICANT: Robinson, Keith E.  
 TITLE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME HOMOLOG  
 NUMBER OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: FOLEY, HONG & ELLIOT LLP  
 STREET: One Post Office Square  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109-2170  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/989,299  
 FILING DATE: 11-DEC-1997  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Arnold E., Beth  
 REGISTRATION NUMBER: 35,430  
 REFERENCE/DOCKET NUMBER: MIA-025.01  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-832-1000  
 TELEFAX: 617-832-7000  
 INFORMATION FOR SEQ ID NO: 6:



RESULT 11  
 US-09-440-325A-1  
 ; Sequence 1, Application US/09440325A  
 ; Patent No. 6280994  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sheppard, Paul O.  
 ; TITLE OF INVENTION: Zaci: A Human Metalloenzyme  
 ; FILE REFERENCE: 98-79  
 ; CURRENT APPLICATION NUMBER: US/09/440.325A  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: FASTSEQ for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 694  
 ; TYPE: PRP  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (1)...(694)  
 ; OTHER INFORMATION: Xaa = Any Amino Acid  
 US-09-440-325A-1

Query Match  
 Best Local Similarity 23.1%; Score 990; DB 4; Length 694;  
 Matches 221; Conservative 104; Mismatches 228; Indels 30; Gaps 15;

22 EFOAKTELDKFNHEAEDLFYOSLSAWNTNTTEENYONMNNAGDKSAFLKQSLAQ 81  
 53 ETEKTFQEDQGEVVLNKFEMATWNTYITTRKNOEMKDMER-SQMIFGTQAH 111  
 82 MPDLOEIONLVKLOALQNGSSVLSDEKSKRLNTILNTMSTYSTGVYCPNDPQEC 141  
 112 LEKTVQFQDPVNGMGLKONIDKALSKDELREYNELLAXLETYSMAOVCLNECP-C 169  
 142 LLEPGNEIMANSIDYNERLAWESWSEVQKOLRPLYEYVYVAKNEMARAH-YEDYG 200  
 170 LSLSEFL-EVMAISRKDELLAMOGMODAVGROICTTFEYVESNKAQOLNGYXKDMG 228  
 201 DYNRGDYEVNGVDYDSRGOLIEDETEETKPLYEHLAAVYRAKIMNAV-PSYISPT 259  
 229 ALMHSKTESDYLE-----QDLRLFOELRPLYLNPHTYVRRALHRYGPELIDLR 278  
 260 GCLPAHLIGD-MMGRFNTNLYSLVPGQKPNIDVTDAVDAQMAORI-FKAEKFEYS 317  
 279 GEPHALLGENTLAQSWNLIDPVLPELKKIPEDVTYKIMKYOHMKPEKMLEEAEFFTY 338  
 318 VG--LPMNTOGFEMENSLDPCVQKAVCHPTAMDIGK-GDFRLMCTKYTMDPFLTAH 374  
 339 LGALPAPPSFWKIMLRPTDGRREVCHISANNFYQODDFRKKCAEVTEDEPLSTFH 398  
 375 EMGIQYDAYAAOPFLIRNGANEGFHAVEITMSAATPKILKSGILSPFOEDNET 434  
 399 EMGHFYFLOYKNLSIIFRTGANPAFEAVGSYTTTASASHKMLNLIGILS--LLED--- 453  
 435 EINFELKQALLTVGLPTTYMLEKRWNVFKGEIPDQMMKRWMEKR-ELIVGVEEYVPH 493  
 454 EVNFMHLIRLEKIAFIPEGYLMDFRKVFVDGIMKDIYDQPMNMLRLKLTQGLCPAIPH 513  
 494 DETYDPAISLHVANDYSFT-RYTRTYLOFOFOBALCOAKHSGPLKHKCDISNSTEAG 552  
 514 SEEPDFDPAKFHSAGVPIYTRFELSLVLOFOFHETLCKSGHNGPLHQCDIYNSKIACK 573  
 553 KLFNMLRIGSEPMVTLALENVGAKNMVNPRLNVPFELTFL 595  
 574 LL--ALKLGSSKFWPEVVLKMLTGESEVSTNVEMTYEKPFLTLWL 614

RESULT 12  
 US-08-989-299-12  
 ; Sequence 12, Application US/08989299  
 ; Patent No. 6194556

GENERAL INFORMATION:  
 APPLICANT: Acton, Susan L.  
 TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: FOLLEY, HOAG & ELLIOT LLP  
 STREET: One Post Office Square  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109-2170  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/989,299  
 FILING DATE: 11-DEC-1997  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Arnold E., Beth  
 REGISTRATION NUMBER: 35,430  
 REFERENCE/DOCKET NUMBER: MIA-025.01  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-832-1000  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 907 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-989-299-12

Query Match  
 Best Local Similarity 15.0%; Score 642.5; DB 4; Length 907;  
 Matches 172; Conservative 136; Mismatches 289; Indels 41; Gaps 19;

2 SSSSLLSLVAVTAAGSTIEQAKFLQKFNHEAEDLFYOSLSAWNTNTTEENYON 61  
 160 SSNMYKTNLOAPGISRD--EKLRSMLAGTEAEKIVLREVALSGRRYTNDSPSKIA 217  
 62 MNNAGDKSAFLKQSLQALQNGSSVLSDEKSKRLNTILN 121  
 218 LDEAEVNLIVFVSTMSQAKQFDMASVDEKVMRQLGYSEFGMSALASRFADYDQA 277  
 122 TWSTYSGVYCPNDPQECILIEPGLNEIMANSIDYNERLAWESWSEVQKOLRPLYE 304  
 278 ALNDSKSTIDKDVPPCALQKIDMSIFRNKAKASRLQHLWASYVAIAKS-KPSYN 336  
 182 EYVVLAKNEMARAHVEDYDGRDYEYVNG-VQGYDSGQLEDETEETKPLYEHL 411  
 337 NITTSNEKAKLNGANGAMMBSAFDMSSKVAKAE---DLNKQIDKTYSTLOPFLQL 493  
 241 HAVYVAKIMNAV-PSYISPTGCLPAHLIGDMKGRFNTNLYSLVPGQKPNIDVTDAV 513  
 394 HAYMRQLAGIYSNPVGLSKDGPFAHLFGSLDGDMSAHYEQKKPFEELS--ETFEAAL 552  
 299 D---QAMDQORIFKAEKFEYSVGLPMNTOGFEMENSLDPCVQKAVCHP-TAMD-LG 570  
 452 SAFNTQYTTKKAFTYATYTRFELSLVLOFOFHETLCKSGHNGPLHQCDIYNSKIACK 595  
 353 KGFRLMCTKYTMDPFLTAHHEMGHIQYDAYAAOPFLIRNGANEGFHAVEITMSA 614  
 511 PNFERYKACAQUGEPDFOASHLVQTYLYTKDOSLLEFQOASPVITDAIANAPHLIS 630  
 413 ATPKHLKSGILSPFOEDNETE-INFLKQALLTVGLPTTYMLEKRWNVFKGEIPD 650  
 571 TNPHTLYSOKLVPESEHLDIKDSYIINKLVKESLSEFTKLPFLTAADMMRYELRPGTVPRN 670

us-09-978-385-2.ra1

```

0Y 472 QMMKKEMKREIVAVPEVPHDEYCDPASLPH--VSNDSYFIRYPTL----XQPF 522
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
631 KLMDRWEIIRNKEVEVSPQPYNTSND--ALHNSQVHS---PATRLISIVLKQI 685
0Y 526 QEALCOA---KHEGPCHKDISNSTAGKLENNMLTCKSEPWLALENYVGNKNNVR 582
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  686 LKALQGRLEFMLSQ---CLISFDYI--EKINETMKLGSSITWLKALEMISGGLDAQ 739
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 583 PLINFEPLFTWLKDQNK--NSFVGWSTQWSPADSI 618
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 740 PLEYEYEPILNIRNTNEIDYVVGMDGETEPTVEI 777

```

US-08-645-193B-15  
Sequence 15, Application US/08645193B  
Patent No. 5962253  
GENERAL INFORMATION:  
APPLICANT: Kupke, Thomas  
APPLICANT: Goltz, Friedrich  
APPLICANT: Kempter, Christoph  
APPLICANT: Jung, Gunther  
TITLE OF INVENTION: Oxidative Decarboxylation of Peptide  
TITLE OF INVENTION: Catalyzed by Flavoprotein EplD  
NUMBER OF SEQUENCES: 70  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/645,193B  
FILING DATE: 13-MAY-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Emond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652.1540000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 990 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
5-08-645-193B-15

Query Match	2.9%;	Score	124;	DB	2;	Length	990;
Best Local Similarity	18.2%;	Pred. NO	0	013.			

```

QY      1  MSSSSWLLSLVAVTAASSTIEQAQKTPIDKRNHE-----AEDL--FYQSSLASWNY 50
Db      324  ISEAAAYILIMLLSPNHGVTIRINYNHEFEDKDGFEQOLVNLKOLLSDINGNGYGRKQSYSP 383
QY      51  NNTNTEENVQNMNMADKMSAFKQSLQSLAQMYPLOEIQNLTVKQLQALQOONGSSVLE 110
Db      384  SNNI-----AFLKPK-----YLLAQNNSHIEITE 408
QY      111  DKSKRLNTLLTMSTIYSTGYKVCNPDPQECLEPGLNEIMANSLDYNRRLAMWSMMS 170
Db      409  NDVKFKLER--NNTVSKINA-----PVSTEL-----YSEIIF----- 437

```

```

QY 171 EVGQJRLPYEEYVLKNEAARANKHEDYGDYRWGDYEVNGVDYDYSRGQILIEVEHTF 230
Db 438 --GMSIKG-YEDFPAVISPIIGSEFAGATGFR-TGNFNIR-----KKNOJKEIYN-- 488
QY 231 EIKPPEYEHLLHAYYRAKIMAAVPSYISPIGCLPAHLHGDPMGREGFTNLXSLYFPGGKPN 230
Db 485 -----HYNNYMNENDLEISQJLNBAPLNSNNVNLNN--NRIYNTCLNINLNP---KSD 531
QY 291 IDVTDAVADQAWPAQRIJFKP---AEKFPSYVOLPMMTGG-----EWSNMLDPC 337
Db 532 IDINDIFIGATFNKILVISEKHEDSRIVAFSNMFYERGESLYFKLEISEFKTKFTOP1 591
QY 338 N-----VQKAVCHPTAMDJGKGFRLIMCTKATMDQFTLAHHEMGHILOYM 383
Db 592 TEECIDSLEPCPRIIYKNTILKPATKINSEMFSE---TENWLRPATI-REKMHIPKDV 647
QY 384 AVA-----AOPLLRNGANGPHEAVGELTMS-----409
Db 648 IIAFGDNRLNLNLNDKHLILKELKKHGRILLESINSENBERLETVLPYLYKTSL 707
QY 410 --LSAATP-----HLK-SIGLSPDFQEDNETEINFLKQALTYIGLPFYMLE 457
Db 708 KEGSFIIPANKRKHNNNLKDMFSTHLSIPKYOD-----NFIDYLLPFTTELKVNNEFN 762
QY 458 KMRVMEYKGEIPKDOMKMKWEMKREIYGVVEYPVPHDETVCBPASLEPHNSNDYSFIRY- 516
Db 763 KEFYIKFKED--EDFIK--LRLLRE-----DEDV-----SQTYSFKMKK 798
QY 517 -----TRLYQFOPEALCOAKHEGHLKCDISNSTEAGOKIENNLRGKSEPITALLE 571
Db 799 DYCLNSELNYDSTVDYEPYRIGTGHVATEDIENEFMYDSL--SINTIOSE-FRIKPE 855
QY 572 NYVGAKNANVAPLNYEPBLFTWMLKDOKNSFVGWSTDMSPYADOSIKVIRISLKSALODK 631
Db 856 FIVA---ISIDFLLDYLE-----INSSEKEILINNA--ED 886
QY 632 AYEMNDNEMETLRSSVAYAMROYFLVKYKNOMLFGSEEDVRANLKRISSENFYTAAPKNV 691
Db 887 LYRSDN-----IREY-----KNLLAKL-----TNPKN 909
QY 692 SDIIP-----RTEVEKAIIRMSRRIINDAF---RLNDSLEFPGIO 728
Db 910 YEILKKEFPNLHEFLNFKISILENLKTLQKSLYTSRSRIIGSFTHMRCN---RINGIN 965
QY 729 P 729
Db 966 P 966

```

US-08-392-625-20  
; Sequence 20, Application US/08392625  
; Patent No. 5837485  
; GENERAL INFORMATION

APPLICANT: Entlan, Karl-Dieter.  
 APPLICANT: G tz, Friedrich.  
 APPLICANT: Schnell, No. 5837485bert  
 APPLICANT: Augustin, Johannes  
 APPLICANT: Engelke, Gernar  
 APPLICANT: Rosensteln, Ralf  
 APPLICANT: Kaletta, Cortina  
 APPLICANT: Klein, Cora  
 APPLICANT: Wieland, Bernd  
 APPLICANT: Kupke, Thomas  
 APPLICANT: Jung, G nther  
 APPLICANT: Kellner, Roland  
 TITLE OF INVENTION: Biosynthetic Process For The Preparation  
 OF SEQUENCES: Of Chemical Compounds  
 NUMBER OF SEQUENCES: 42  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Steme, Kessler, Goldstein & Fox  
 STREET: 1100 New York Avenue

Fri Mar 14 10:00:48 2003

us-09-978-385-2.ra1

Page 10

CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/392,625  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/876,791  
FILING DATE: 30-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652.0980002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 990 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-392-625-20

Query Match 2.8%; Score 120; DB 2; Length 990;  
Best Local Similarity 17.8%; Pred. No. 0.031;  
Matches 152; Conservative 121; Mismatches 246; Indels 334; Gaps 41;

QY 1 MSSSWLLSLVAVTAOSTIEQAKTPIDKNHE-----AEDL--FYOSSIASMNY 50  
DB 324 ISEAAVTLMLSPHEFGTITRNYHEFMDKGFQOLYNLQSLDINGFPYKDDYSF 383  
QY 51 NTNTEENVQNMNAGDKSAFLKEQSTLAQWPIQETQNTLVKIQALQONSSVLS 110  
DB 384 SNNI-----AFLEK-----YLLAIQNNSHIETTE 408  
QY 111 DKSRRLTINTMSTIYSTGVKCPNPDPQCLLEPGLEIMANSLDNERLMAESWRS 170  
DB 409 NDVAKLEK--NNTVSKINA-----PVSTET-----YSEIIF----- 437  
QY 171 EVGKQRLPELYEYVLKKNEMARAHYEDYGDYWRGDEYVNGVDYDYSRGQLIEDVET 230  
DB 438 --GNSIKG-YEDFAVISPIIGSFAGATGFRP--TGNEFNK-----KKNOLQKEIVH-- 484  
QY 231 BEIKPLVHLLAVYRAKLMNAYPSYISPIGCLPAHLLGDMGWGFEWNTILSLVPPGOKPN 290  
DB 485 -----HNNYMNNDLEISQINPAFLNSRVNLTINN--NRITYTCLNLMP--KSD 531  
QY 291 IDVPMADVQAMDQRIKPE---AEKFEVSGVLPNMTQG-----FWNSMLTDPG 337  
DB 532 IDINDIFGATFNKLIYSEKHDRIVFVSNMNYEFGSELKFPREISFEKTRIOPI 591  
QY 338 N-----VQAVCHPTAWDLGK-----G 354  
DB 592 TEGCIDSLPFCRIIYKNIILKPTWKINSEMESETENMLNRAVITRKWIIPKDYIANG 651  
QY 355 DFIILMCTVMTDDELTAHH-----EMGHIOYDMAYAAOPELLRNGANGEHVANG 405  
DB 652 DNLLL-----NLNDKHLIILKELKHGRIR-----LIEPSINESNBERMI 694  
QY 406 EIMS-----LSAATPK-----HLK-----SIGLSPPQEDMETELNFLQAL 445  
DB 695 EIVTPLYKKYSLEKQSFIIPIKRNKHFNNLKDWFESIHLSIKTYOD--NFLQDYLLP 749  
QY 446 IVGTLPPTTYLEKWRMVAKEGELPKDQMKKWEKREIVGVVPHDETQCDPASLFH 505  
DB 750 FTELKVNNEFNKFFYIKRED--EDFIK--LRLIRE-----DEDY----- 786

QY 506 VSNDSYFIRY-----TRLLQFOFOEALCOAKHEGPHKCDISNSTEAGOKLENNLR 559  
DB 787 -SQIYSFIKNWKDYCLINSELXYDSYDVEYRYRGGHVIEDLENF--MDSLDSDIN 844  
QY 560 LCKSEPTLALENVYAKKNMVRPLNYPFLPTFWLKDONKNSFVGWSTWMSPYDOSIK 619  
DB 845 IIQSE-FKIPKEFYA--ISIDPLDYLE-----INSEK 876  
QY 620 VRISSKALGKAYEWNDEMYLFRSSVAYAMQYFLAKYKNOMILFGEEDVAVANLKPRI 679  
DB 877 EELLNNA--EDLYRSND-----IRY-----KNILAKI 903  
QY 680 SFNFVYTAKNVSDIIP-----RTEVKAIRMSRINDAF--R 716  
DB 904 -----TNPKNRYEILKKEFPLHIEFLFNKISILENKLKTLQSLVTSRSRIIGSFIMHR 957  
QY 717 LNDNSLEFLGIOP 729  
DB 958 CN-----RIFGINP 966

RESULT 15  
US-08-466-961A-20  
Sequence 20, Application US/08466961A  
Patent No. 5843709  
GENERAL INFORMATION:  
APPLICANT: Entian, Karl-Dieter  
APPLICANT: G tz, Friedrich  
APPLICANT: Schnell, No. 5843709bert  
APPLICANT: Augustin, Johannes  
APPLICANT: Engelke, Gerhard  
APPLICANT: Rosenstein, Ralf  
APPLICANT: Kaletta, Corlina  
APPLICANT: Klein, Cora  
APPLICANT: Wieland, Bernd  
APPLICANT: Kupke, Thomas  
APPLICANT: Jung, G nther  
APPLICANT: Kellner, Roland  
TITLE OF INVENTION: Biosynthetic Process for the Preparation of  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, NW  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,961A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/392,625  
FILING DATE: 22-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/876,791  
FILING DATE: 30-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/784,234  
FILING DATE: 31-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 0652.0980004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600

TELEFAX: (202) 371-2540  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 990 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 US-08-466-961A-20

Query Match 2.8%; Score 120; DB 2; Length 990;  
 Best Local Similarity 17.8%; Pred. No. 0.031;  
 Matches 152; Conservative 121; Mismatches 246; Indels 334; Gaps 41;

```

QY 1 MSSSSWLLSLVAVYAAQSTEEBAKTFLEKFNHE-----AEDL--FYQSSLASMY 50
DB 324 ISEAAYITLMLSPHEFGTKITRYNHEFEFMDKGFQEDLVNMLKQLLSDINGFCYPKKDSIF 383
QY 51 NTNTEENVQNMNAGKWSAFLKEQSTLAOMVPIQETQNLVKIQALQALQNGSSVLYSE 110
DB 384 SNNI-----AFLEK-----YLLAIQNNSHIRETE 408
QY 111 DSKRLTILNTMSTYTGKVCNPNPOECLELPEGLNEIMANSIDYNERIAMESWRS 170
DB 409 NDVKNLEK--NTVSKINA-----PVSTET-----YSELYF----- 437
QY 171 EVCKORPLVEEYVLKEMARANYEDYGDYWRGDYEVNVDGYDYSQGLIEDYEHTE 230
DB 438 --GNSIKG--YEDFAVISPLIGSFNAGATGRRP--TGNFNK-----KKQLOKEIVH-- 484
QY 231 EETKPLYEHLAAVYRAKLNNVPSYISPGCLPAHLGDMWGRFWTNVSLVTPGOKRN 290
DB 485 -----HYNNMENDLEISQLEAPLNSRVNNILNN--NRITYTCLNLMP---KSD 531
QY 291 IYVTAMDQADQRIKFE---AEKFFVSGLEPNTOG-----FWNSMLTDPG 337
DB 532 IDINDIEIGAFNKLYLSEKHDSRIYFVSNMFNFESESELYKELREISEKTFIQPI 591
QY 338 N-----VOKAVOHPAMDGK-----G 354
DB 592 TEEGIDSLPFCPRITLYKNIILKPAIWKINSEMESETEMLNFRATIRKWHIPKVIITAFG 651
QY 355 DFRIMGCKTYMDFLTAH-----EMGHIOYDAVAAQAPPLANGNEGFHVAVG 405
DB 652 DNRILL-----NLNDKHLILKELKELKHGRIR-----LLESFINESNMERM 694
QY 406 EIMS-----LSAATPK-----HLK---SIGLSPDQEDNETEINFLLQALH 445
DB 695 EIVTPLYKRTSLKEQSFIPKRNKHNKLDKMFSLHSIPKTYQD-----NFIOYLLP 749
QY 446 IYGLPPTYLEKRWMPFKGEIPKQDMKMKWMEKREIYGVYEPVPHDTEYDPAFLPH 505
DB 750 FTELAVNNFNKFFYIKKED--EDFIK--LRLLE-----DEDY----- 786
QY 506 VSNDSYFIRY-----TRTYFOFOALCOAAKHGRLHKCDISNSTEAGOKLFNMR 559
DB 787 -SOIYSFINKWKDCLLSELXYDIYDYVPEYRYRGQGVHIDIDNEF--MDSLIDSLIN 844
QY 560 LKSEPMWLALENVYGAKMNVRLPLNYFEPLFTWLKDKNKSFWQSTMSDVSQDQK 619
DB 845 IIGSE-FRIPEFIVA---ISIDFLDYLE-----INKSEK 876
QY 620 VHSLSKSLGDKAYENDNEMVFLRSSVAVAMQVFLAKVKNOMILFGEDEVAVANLKPRI 679
DB 877 EELINNA--EDLYRSND-----IREY-----KNLAKL 903
QY 680 SFNEFVAPKRVSDIIP-----RTEVEKAIKMSRSFINDAF---R 716
DB 904 -----INPNANDYELIKKEFPNLHEFLFNKISILENLKKTLOKSLYTSRSHIGSFTHMR 957
QY 717 LNDNSLEFLGIOP 729
DB 958 CN-----RIEGINP 966
  
```

## RESULT 16

US-08-630-916A-46

Sequence 46; Application US/08630916A

Patent No. 6011137

GENERAL INFORMATION:

APPLICANT: Pirozzi, Gregorio

APPLICANT: Kay, Brian K.

TITLE OF INVENTION: Fowlkes, Dana M.

TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL

NUMBER OF SEQUENCES: 124

CORRESPONDENCE ADDRESS:

ADDRESSER: Pennie &amp; Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: United States

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,916A

FILING DATE: 03-Apr-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MISOCK, S. LESLIE

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-203

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 896-8664/9741

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 683 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: Peptide

US-08-630-916A-46

## Query Match

Best Local Similarity 2.7%; Score 116.5; DB 3; Length 683;

Matches 142; Conservative 110; Mismatches 268; Indels 269; Gaps 36;

```

QY 12 VAVTAQSTIEQAKTFLEKFNHEAEDLFYQSSLASMYNTNTEERNVQNMNAGDKWSA 71
DB 44 LASPADDTVNGESSFAPTN-----ASVTGPVVSSEMLNSPCT----- 85
QY 72 FLKEQSTLAQMTPIQET-----QNTLVKIQALQALQNGSSVLYSEKSKRLTILNTMSTI 126
DB 86 -----STVEDPVOQELITSSENNCEIPSTSAELSEASLIEPRTSN-----SRSSA 134
QY 127 YSTGVCNPNPOECLELPEGLNEIMANS-----LDYNERIAMWS 167
DB 135 FEAASKRQPDGCMQDPVROQSG--NANTETLPSGWEQRKDPHGRITTYVVDHNRITTWER 190
QY 168 WRSEVGKQLRPLIEEYVLKEMARANYEDYGDYWRGDYEVNVDGYDYSQGLIEDY 226
DB 191 -----POPLPPEGRVVDORRRYVVDHNRITTWERQRTMESVAVNFQWQOSORNOI----- 241
QY 227 EHTFEIKPLYEHLAAVYRAKLNNVPSYISPGCLPAHLGDMWGRFWTNVSLVTPFG 286
DB 242 QGAMQOFNORYLYSASMLAAB-----NDPTGPLP-----G 272
QY 287 OKPNIDVTAM-----VDQAMDAQRIKFEKAEFVSGLP 321
DB 273 WEKRVDSITRVYFVHNKTTQWEDPRTQGLQNEEPLPEGWEI--RTYRAGVYFVDHNTNR 331
QY 322 NMTQGFWNSMLTDPGNOVAKVCH---PTAMDIGKDKDFRLKCTKVYMDFLTAHENGH 378
  
```



Fri Mar 14 10:00:48 2003

us-09-978-385-2.ral

Db 332 TTT-----FKDPRNGKSSVTKGPOJAYERG---FR-----KLAH 364  
OY 379 IQYMAYAAOPELLRNGANGCEHVAEGEIMSLSAATPRHLKSLGLSPDOEDNEIEIN 438  
Db 365 FRY-----LCOSNMLPSHVK--INVSROTLFEDESFQOI-M 396  
OY 439 LKQALITVGLPTTYMLEKRMWFKGELIPKQO--MMKKWEMKREIVGVVEV----- 491  
Db 397 ALK-----PYDLRRLRYIFRGEGLDYGGLAEWFFLLSH--EVLNPRYCIFE 443  
OY 492 -PHDETC--DPASLF--HVSNDYSFI-RVYTRLYOFQOEALCOAHNEGPLK-- 541  
Db 444 YAGKNNYCLOINPASTINPDLSY-FCFIFGRITAMALFNGKFDGFSLPFKRLSKL 502  
OY 542 --CDI-----SNSTEGOKLF--NMRLKSKSEPTTALLENVYGAKNMN 580  
Db 503 TIKDLISIDTEFYNSLIWIRONNIEGCELEKTFVDMKELIKVTSIDLK---GGSNLT 558  
OY 581 VRPLWVFEPLFTWLKDKNSFGWSTDSVPADOSIKYRISLKSALDQKAYEMNDNM 640  
Db 559 V-----TEENKDEYIGLMTW-----RFSKQVOEQTKAFIDGPFNEV 594  
OY 641 YLRSVAYAMROYFLKVNOMILFGEEDVRVANKPRISENFEVTAAPKNVS---DIIFR 697  
Db 595 -----VPLQMLQYFDEKELEWALCGMQEVDLADMQRNTVYRHYTRNSKOIIMWQFVKE 648  
OY 698 TEVEKAIKM 706  
Db 649 TDNEVBMRL 657

RESULT 17  
US-08-844-059-2  
; Sequence 2, Application US/08844059

GENERAL INFORMATION:  
PATENT No. 6001601  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6001601el Compounds  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19046  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/844,059  
FILING DATE: 18-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION NUMBER: 9607999.1  
FILING DATE: 16-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Gimmil, Edward R  
REGISTRATION NUMBER: 38,891  
REFERENCE/DOCKET NUMBER: P31456  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-4478  
TELEFAX: 610-270-5090  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 665 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-844-059-2

Query Match 2.6%; Score 113.5; DB 3; Length 665;  
Best Local Similarity 19.2%; Pred. No. 0.067;  
Matches 135; Conservative 99; Mismatches 231; Indels 239; Gaps 33;

OY 143 LLEPGLNEIMANSLDYNELNAMESRSEVQKOLRLPYEEVVLKKNMARMANHYEDYD 202  
Db 85 LDISYDKFRTTDDYHEKVA-----QVEERLLAODDIY--LGEY 123  
OY 203 WRGDEVNGVDGYDSNGOLIE-----DVETFEETRLPYEHLNAVY 244  
Db 124 -SGWTSYS--DEEFTFESQALAEVFRDEAGNVYGIAPSGHEVWSESTFELKSKYQR 180  
OY 245 RAKIMNAVPSYISYIGLPAHL-----LGMMSRFTVNTYSLTVPGOKPNIDV-TDA 296  
Db 181 LVEEFKAPHEPITPDGRLENLKNFIEPGLIEDLAASRTFTWGVVPSNPKHVVYVWIDA 240  
OY 297 NV-----DOAMDA--QRIKFAEKF-----FVSGLPN--MTQ 325  
Db 241 LNTATALGVAODEHGNDFKFMNGTVFHWGKIDLFHSIYPIILMLMDYKLDRLIAH 300  
OY 326 GFV--ENSLMTPDGNVQKAVCHPTAMDLGKGDRIIMCTKYTMDDELTAHHEMGHIOYD 382  
Db 301 GWFVMDGMSKSKGNVYPRMILVERYGDLRLYILRNLPVGSDFTEPDEYVGRIVE 360  
OY 383 MAYAAOPELLR-----NGANEGHEAVGEIMSLSAATPRHLKSLGLSPDOEDNE 433  
Db 361 LANDLGNLNRVTVSMINKYFDQIPAYEGVTEPDHVALE--KSLA---DEFTIME 414  
OY 434 TEINELLKQALITVGLT--PPTYMLEKRMWFKGELIPKQO--MMKKWEMKREIVGV 487  
Db 415 A-VDY--PRALLEVWTLISRTNKYIDETAPVVLKDEALDQSLASVSHQASIRVVAHL 471  
OY 488 VEPVHDEYCDPASLFHVSNDYSFIRYTRTYIOFOEALCOAHNEGPLKCDISNS 547  
Db 472 IEP-----FMETSRVY----- 483  
OY 548 TEAQOKLFNMLRLKSEPTTALLENV-----VAKNMNVRPLNYEPLFTWLKD 597  
Db 484 -----LQIGLEE--VSSLENLSLADFPADVTVVAKGTPIFPRLNMEEL-AIKRE 531  
OY 598 QKNSFGWSTDSNP-----YAD-QSIRKVRIS-----LKSALDQKAYE-- 635  
Db 532 QMEGNKPAVEKEWMEDEVELKINKDELKFEDEPKVIEIRVAEVEKSVGSGDKLLQRLD 591  
OY 636 -NDNEKYLFRSSVAYAMROYFLKVNOMILFGEEDVRVANKPRISFNFVTAAPKNVSDI 694  
Db 592 AGDGEDROLISGIA---KYV--PNOELVGRKVOIVANKFR----- 628  
OY 695 IPRTVEKAIKRSRINDAFRL--NDNSLFEFLGIOTPLGPPN 735  
Db 629 -----KMKKRYVSGAMLSAHDGKRLTLTVDPVAV--PN 660

RESULT 18  
US-09-431-202-2  
; Sequence 2, Application US/09431202

GENERAL INFORMATION:  
PATENT No. 6294175  
APPLICANT: Lawlor, Elizabeth  
TITLE OF INVENTION: No. 6294175el Compounds  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Smithkline Beecham Corporation  
STREET: 709 Swedeland Road  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19046  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/431,202  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/844,059  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gimm, Edward R  
 REGISTRATION NUMBER: 38,891  
 REFERENCE/DOCKET NUMBER: P31456  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 610-270-4478  
 TELEFAX: 610-270-5090  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 665 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-431-202-2

Query Match 2.6%; Score 113.5; DB 4; Length 665;  
 Best Local Similarity 19.2%; Pred. No. 0.067;  
 Matches 135; Conservative 99; Mismatches 231; Indels 239; Gaps 33;

13 LLEPLNIMANSIDYNERLMAWMSRSEVKGOLRPLYEEVYLKEMARAHYEDYGY 202  
 85 LLDISYDEFIKTTDYHEKVA-----QVEERLADDDIY--LGEX 123  
 203 WRGDYEVNGVDYSRGOLIE-----DVEHTTEELKPLYEHAAV 244  
 124 -SGWYSVS--DEEFTESSQLAEVFRDAGVNTGSIAPSGHEVWSEESTYRLSKYODR 180  
 245 RAKLMAVPTISPTGCLPAHL-----LGMWGRFTNLYSLVYFQCKPNIIDY--TDA 296  
 181 LVFEFKAHPEITPDGRLEMLNENFIEPGEDLAVSRFTFTGVPVPSPKHVVYWDIA 240  
 297 MV-----DOAWDA--QRIKFAEF-----FVSGLPN--MTQ 325  
 241 LLNATATLGAODEHGFDFEFMGVYFHMKGDLRFHSIYWPILMLMDVRLPRLIAH 300  
 326 GFV---ENSMILDPGNVQAKVCHPTAMDIGKDFRILMCTKYTMDFLAHEMHQIYD 382  
 301 GWFYMKDKMSKSGNVVPEMLVERYGIDPLRYILMRMLPVGSDGTFTPEDEYGRINTE 360  
 383 MAYAAGPEFLR-----NGANEGFHEAVGEIMSLSATPKHUKSIGLLSPDQEDNE 433  
 361 LANDLGNLNRVSMINKYFDGIPAYVEGVTEFDHVLAEVAE--KSLA---DEHTHME 414  
 434 TEINFLKQALITVTL---PFTYMLEKRWMMVFKEIPKQDM---MKKWMKKEIYGV 487  
 415 A--DY--PRALAEVWTLIRSTNKYIDETAPWLDKDEALRQDLASVMSHQAISIVAH 471  
 488 VEEVPHDETCDPASLFHVSNDYSFRTYRTLYQFQEQEALCOAAKHEGPLKCDISNS 547  
 472 IEP-----FMETSRVAV----- 483  
 548 TEAGOKLFNMLRLGKSEPTLALENV-----VGAKNNVNPRLNTEPEPLTWLKD 597  
 484 -----LTOIGLEE--VSSLENISLADPADVTVVAKGPTIPRLNMEETI--AYIKI 531  
 598 QAKNSVSGSTWSP-----YAD--QSIKVRIS-----LKSALDKKAYEV--- 635  
 532 QMGKNPAVEKEMNDPEVLEKLKDEIKFEDEDKVEIRHVAEYKESVYSGSKDLQFRLD 591  
 636 -NDNEMYLFRSSVAYAMROYELKVNOMILREEDVYRANLKPRTSFEVYTAAPKNSDI 694  
 592 ADGGEKROLISGIA-----KTY--PNEOELVGKKVQIVANLKP----- 628

QY 695 IPTEVEKAIKMSKRNINDAFRL---NDNSLEFLGLOPTLGPNN 735  
 DB 629 -----KMKKVVSGMILSAHDKLTLITLVDRVAV--PN 660

## RESULT 19

US-08-491-357-2  
 Sequence 2, Application US/08491357  
 Patent No. 5716782  
 GENERAL INFORMATION:  
 APPLICANT: Coleman, Erica A.  
 APPLICANT: Law, Susan  
 APPLICANT: Estojak, JoAnne  
 TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING A SIGNAL  
 TITLE OF INVENTION: MEDIATOR PROTEIN THAT INDUCES CELLULAR MORPHOLOGICAL  
 TITLE OF INVENTION: ALTERATIONS  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Dann, Dorfman, Herrell and Skillman  
 STREET: 1601 Market Street Suite 720  
 CITY: Philadelphia  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19103-2307  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/491,357  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Reed, Janet E.  
 REGISTRATION NUMBER: 36,252  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (215) 563-4100  
 TELEFAX: (215) 563-4044  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 834 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: protein  
 HYPOTHEtical: NO  
 ANTI-SENSE: NO  
 US-08-491-357-2

Query Match 2.4%; Score 104; DB 1; Length 834;  
 Best Local Similarity 20.5%; Pred. No. 0.81;  
 Matches 88; Conservative 59; Mismatches 142; Indels 140; Gaps 19;

QY 56 EENVQNNMAGDKWSAFLE-----OSTLAQWYPLQE--IONITVYLOLOALQONSSVLS 109  
 DB 442 ERHINEIRAVKVEFLFLEBYLHFYKGAVANACLPDLHKMKRREIQRV--EDSHQILS 500  
 QY 110 E-----DKSRKRLTILNT----- 122  
 DB 501 QTSHDNCSWSLNTILAIKPNKCDLDRFYVAVAKTVVDDAKQLTITNTINAEALFRPG 560  
 QY 123 -----NSTY-----STGKYCNP--DNQEC--LLLEPLGNETMA---NSLDY 158  
 DB 561 PGSLHAKNPESIMNSTEYPRHGGSGOLHPEHDKRAQAHNRALPPLGSLKEQADPCSSSDG 620  
 QY 159 NERLMAWE-----SWRSEVGRQLRPLYEEVYVLKNEVARAHYEDYGYWRGDYEVNGV 212  
 DB 621 SERSWDDYDYVHLQGEKEFEHQKLEKEMIKNNKMLEHHQ----- 665  
 QY 213 DGVDYSRQGLIEDVHTTEELKPLYEHAAVYRAKLMANVPSYISPTGCLPAHLIGDMWG 272

```

Db 666 ----LSQFOLLBQ-----EITKPVENDISKW---KPSQSLPTTNSGVSAODRQLLCFYVD 713
QY 273 RFWNTNLSLTFVPGCKPNIDVTDAWVDAQMAQO--RIFKEAEKF-----FVSGGLPNM 323
Db 714 QCFTHFTSL-----LNALDALFSCVSSAOPRITVAHSKYVILSAHKLTVIG-DTL 763
QY 324 TQGFWENSMLTDPGNVQAKVCHPTAWDLGKGDGDFRILMCTKYTMDF--LTAHHEMGHIQY 381
Db 764 TRQVTAODIRNKVMNNSNQLCEQLK-----TIYMATKMAALHPSTTALQEWVHQVT 815
QY 382 DMAYAAQPF 390
Db 816 DLSRNAQLF 824

RESULT 20
US-08-968-633-2
; Sequence 2, Application US/08968633
; Patent No. 6100384
; GENERAL INFORMATION:
; APPLICANT: Golemis, Erica A.
; APPLICANT: Law, Susan
; APPLICANT: Estojak, Joanne
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING A SIGNAL
; TITLE OF INVENTION: MEDIATOR PROTEIN THAT INDUCES CELLULAR MORPHOLOGICAL
; TITLE OF INVENTION: ALTERATIONS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street Suite 720
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,633
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/491,357
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Reed, Janet E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; STRANDEDNESS: NO, 6100384 Relevant
; TOPOLOGY: NO, 6100384 Relevant
; MOLECULE TYPE: Protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-968-633-2

Query Match 2.4%; Score 104; DB 3; Length 834;
Best Local Similarity 20.5%; Pred. No. 0.81;
Matches 88; Conservative 59; Mismatches 142; Indels 140; Gaps 19;

QY 56 EENVYOMNNAAGDKWSAFLEKE-----QSTLAQMYPLQE-TONTLVKLOLQALQNGSSVLS 109
Db 442 EKHINEIRTAVDKVELFLKEVYHFWKAVANNAACLPILLINKMKREIQRY-EDSHQILS 500
QY 110 E-----DKSKRLNTLTNT----- 122

```

```

Db 501 QTSHDLNECSMNLNLAINKPNKCDLDRFVNVAKTVPDDAKQLTTINTNAEALFRPG 560
QY 123 -----MSTIY-----STGKVCNP-DNPQEC--LLEPPGUNEIMA---NSLDY 158
Db 561 PGSLHLKNGPESIMNSTEYPHGSOGOLHHPGDHKAQAHNKAALPPGLSKEDAPDCSSDG 620
QY 159 NERLNAME-----SWRSEVGRQLRPLYEYVVLKNEMARANHEDEDGWDYRGDYEYNGV 212
Db 621 SERSMWDDYDVYHLOGKEFEFEKQKELLEKENIKQNKQMLEHHQ----- 665
QY 213 DGYDYSRQGLIDYVHTFEETKPLYEHLHAYVRAKLMNAYPSYISPGCLPAHLGDMRG 272
Db 666 ----LSQFOLLBQ-----EITKPVENDISKW---KPSQSLPTTNSGVSAODRQLLCFYVD 713
QY 273 RFWNTNLSLTFVPGCKPNIDVTDAWVDAQMAQO--RIFKEAEKF-----FVSGGLPNM 323
Db 714 QCFTHFTSL-----LNALDALFSCVSSAOPRITVAHSKYVILSAHKLTVIG-DTL 763
QY 324 TQGFWENSMLTDPGNVQAKVCHPTAWDLGKGDGDFRILMCTKYTMDF--LTAHHEMGHIQY 381
Db 764 TRQVTAODIRNKVMNNSNQLCEQLK-----TIYMATKMAALHPSTTALQEWVHQVT 815
QY 382 DMAYAAQPF 390
Db 816 DLSRNAQLF 824

```

Search completed: March 13, 2003, 16:58:30  
 Job time : 29 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2003, 16:58:07 ; Search time 17 Seconds  
(without alignments)  
2182.595 Million cell updates/sec

Title: US-09-978-385-2

Perfect score: 4291  
Sequence: 1 MSSSSWLLSLVAVTAQST.....ISKGNPNPQNTDQYTSF 805

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PTCT\_NEW\_PUB pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/PTCTUS\_PUBCOMB pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4291	100.0	805	US-09-978-385-2	Sequence 2, Appl1
2	4287	99.9	805	US-10-114-893-86	Sequence 86, Appl1
3	3775	88.0	711	US-09-969-384-13	Sequence 13, Appl1
4	3643	84.9	681	US-09-969-384-25	Sequence 25, Appl1
5	3579	83.4	805	US-09-978-385-6	Sequence 6, Appl1
6	3561	83.0	805	US-09-978-385-9	Sequence 9, Appl1
7	2879	69.4	555	US-10-028-072-72	Sequence 72, Appl1
8	2879	69.4	555	US-10-121-049-72	Sequence 72, Appl1
9	2879	69.4	555	US-10-123-904-72	Sequence 72, Appl1
10	2879	69.4	555	US-10-140-470-72	Sequence 72, Appl1
11	2879	69.4	555	US-10-175-746-72	Sequence 72, Appl1
12	2879	69.4	555	US-10-176-918-72	Sequence 72, Appl1
13	2879	69.4	555	US-10-176-921-72	Sequence 72, Appl1
14	2879	69.4	555	US-10-137-865-72	Sequence 72, Appl1
15	2879	69.4	555	US-10-140-474-72	Sequence 72, Appl1
16	2879	69.4	555	US-10-142-431-72	Sequence 72, Appl1
17	2879	69.4	555	US-10-143-114-72	Sequence 72, Appl1
18	2879	69.4	555	US-10-140-002-72	Sequence 72, Appl1
19	2879	69.4	555	US-10-142-419-72	Sequence 72, Appl1

20	1359	31.7	261	9	US-09-969-384-23	Sequence 23, Appl1
21	1337	31.2	1265	9	US-09-964-899-19	Sequence 19, Appl1
22	990	23.1	694	10	US-09-846-996A-1	Sequence 1, Appl1
23	376	8.8	212	9	US-09-992-598-387	Sequence 387, App
24	376	8.8	212	9	US-09-889-293A-387	Sequence 387, App
25	376	8.8	212	9	US-09-889-735-387	Sequence 387, App
26	376	8.8	212	9	US-09-990-444-387	Sequence 387, App
27	376	8.8	212	9	US-09-989-730-387	Sequence 387, App
28	376	8.8	212	9	US-09-990-436-387	Sequence 387, App
29	376	8.8	212	9	US-09-991-181-387	Sequence 387, App
30	376	8.8	212	9	US-09-893-687-387	Sequence 387, App
31	376	8.8	212	9	US-09-889-734-387	Sequence 387, App
32	376	8.8	212	9	US-10-028-072-482	Sequence 482, App
33	376	8.8	212	9	US-09-997-653-387	Sequence 387, App
34	376	8.8	212	9	US-09-993-667-387	Sequence 387, App
35	376	8.8	212	9	US-10-121-049-482	Sequence 482, App
36	376	8.8	212	9	US-10-123-904-482	Sequence 482, App
37	376	8.8	212	9	US-10-140-470-482	Sequence 482, App
38	376	8.8	212	9	US-09-990-438-387	Sequence 387, App
39	376	8.8	212	9	US-09-990-562-387	Sequence 387, App
40	376	8.8	212	9	US-09-997-428-387	Sequence 387, App
41	376	8.8	212	9	US-09-997-666-387	Sequence 387, App
42	376	8.8	212	9	US-10-175-746-482	Sequence 482, App
43	376	8.8	212	9	US-10-176-918-482	Sequence 482, App
44	376	8.8	212	9	US-10-176-921-482	Sequence 482, App
45	376	8.8	212	9	US-09-990-711-387	Sequence 387, App

## ALIGNMENTS

RESULT 1  
US-09-978-385-2  
Sequence 2, Application US/09978385  
Patient No. US20020177211A1  
GENERAL INFORMATION:  
APPLICANT: Piddington, Christopher S.  
APPLICANT: Petrie, Charles  
APPLICANT: Shoemaker, Kimberly E.  
APPLICANT: Bishop, Paul D.  
TITLE OF INVENTION: ZAC2: A HUMAN METALLOENZYME  
FILE REFERENCE: 99-24CI  
CURRENT APPLICATION NUMBER: US/09/978,385.  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 60/133,952  
PRIOR FILING DATE: 1999-05-13  
PRIOR APPLICATION NUMBER: 60/151,181  
PRIOR FILING DATE: 1999-08-27  
PRIOR APPLICATION NUMBER: 09/563,516  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 805  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-978-385-2  
Query Match 100.0%; Score 4291; DB 9; Length 805;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 805; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSSSSWLLSLVAVTAQSTIEQAKFTLDKFNHEADLFYQSSLSAMWNTNTEENVO 60  
DB 1 MSSSSWLLSLVAVTAQSTIEQAKFTLDKFNHEADLFYQSSLSAMWNTNTEENVO 60  
QY 61 MNMNAQKWSAFLKEOSTLAQNTPLQEIQNLTVKQLQALQNGSSVLSDEKSKRLNTLL 120  
DB 61 MNMNAQKWSAFLKEOSTLAQNTPLQEIQNLTVKQLQALQNGSSVLSDEKSKRLNTLL 120  
QY 121 NTMSTIYSTGKVCNPDNDPQCELLLEPGLNIMANSIDYNERLWAMSWSEVQKOLRPLX 180  
DB 121 NTMSTIYSTGKVCNPDNDPQCELLLEPGLNIMANSIDYNERLWAMSWSEVQKOLRPLX 180

```

QY 181 EBYVVLKNEARAHNYEDYGDYWRGDEYVNGVDGYDSRGQLIEDVHTFEIEIKPLYEHL 240
D 181 EBYVVLKNEARAHNYEDYGDYWRGDEYVNGVDGYDSRGQLIEDVHTFEIEIKPLYEHL 240
QY 241 HAYYRAKLNNAPSYISPIGCLPAHLILGDMGGRFTNLXSLTVPFGKPNIDVTDAWDQ 300
D 241 HAYYRAKLNNAPSYISPIGCLPAHLILGDMGGRFTNLXSLTVPFGKPNIDVTDAWDQ 300
QY 301 AMDAORIFKEAEKFFVSVGLPNNMTOGFWNSMLTDPGNVQKAVCHPAMDLGKDFRILM 360
D 301 AMDAORIFKEAEKFFVSVGLPNNMTOGFWNSMLTDPGNVQKAVCHPAMDLGKDFRILM 360
QY 361 CTKYTMDFLTAAHHEMGIQYDMAAOPFLLRNANGEGHEAVGEIMLSAATPKHLKS 420
D 361 CTKYTMDFLTAAHHEMGIQYDMAAOPFLLRNANGEGHEAVGEIMLSAATPKHLKS 420
QY 421 IGLSPDFOEDNETEINFLKQALTYGTLPFTYMLEKRWMPFKGEIPKQDMKKWEM 480
D 421 IGLSPDFOEDNETEINFLKQALTYGTLPFTYMLEKRWMPFKGEIPKQDMKKWEM 480
QY 481 KREIVGVPEVPDHECYCDPASLFHVSNDSFIRYTRTLXOFQOALCOAAKHGEPHL 540
D 481 KREIVGVPEVPDHECYCDPASLFHVSNDSFIRYTRTLXOFQOALCOAAKHGEPHL 540
QY 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVYGAKNMVRPLNFEPLFTWLKDQNK 600
D 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVYGAKNMVRPLNFEPLFTWLKDQNK 600
QY 601 NSFVGMSTDMSPYADOSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVN 660
D 601 NSFVGMSTDMSPYADOSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVN 660
QY 661 QMILFGEEDVRANLKRISFNFEVYAPKAVSDIIPTEVEKAIKRSRSRINDAFRLND 720
D 661 QMILFGEEDVRANLKRISFNFEVYAPKAVSDIIPTEVEKAIKRSRSRINDAFRLND 720
QY 721 SLEFLGIQPTLGPDPNPVSIMLIVGVVGVIVGIVILFTGIRDKKKKARSGENP 780
D 721 SLEFLGIQPTLGPDPNPVSIMLIVGVVGVIVGIVILFTGIRDKKKKARSGENP 780
QY 781 YASIDISKGNNPGFQNTDVOYTSF 805
D 781 YASIDISKGNNPGFQNTDVOYTSF 805

RESULT 2
US-10-114-893-86
; Sequence 86, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavalite, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, Mckeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; EARLIER FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,332
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 86
; LENGTH: 805

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-893-86
Query Match 99.9%; Score 4287; DB 9; Length 805;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 804; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSWLLSLVAATAOSTIEQAKTFPLDKRNHAEDELFOSSSLASNNYNTITEENVO 60
D 1 MSSSWLLSLVAATAOSTIEQAKTFPLDKRNHAEDELFOSSSLASNNYNTITEENVO 60
QY 61 NNNAGDKWSAEFLKQSTLAQWYPIQETQNLTVKLOLQALQNGSSVSEDKSKRLNTIL 120
D 61 NNNAGDKWSAEFLKQSTLAQWYPIQETQNLTVKLOLQALQNGSSVSEDKSKRLNTIL 120
QY 121 NTMSTIYSTGKVCNDNPQECILLEPGLNEIMANSLDYNERLAMESRSEYQOLRPLY 180
D 121 NTMSTIYSTGKVCNDNPQECILLEPGLNEIMANSLDYNERLAMESRSEYQOLRPLY 180
QY 181 EBYVVLKNEARAHNYEDYGDYWRGDEYVNGVDGYDSRGQLIEDVHTFEIEIKPLYEHL 240
D 181 EBYVVLKNEARAHNYEDYGDYWRGDEYVNGVDGYDSRGQLIEDVHTFEIEIKPLYEHL 240
QY 241 HAYYRAKLNNAPSYISPIGCLPAHLILGDMGGRFTNLXSLTVPFGKPNIDVTDAWDQ 300
D 241 HAYYRAKLNNAPSYISPIGCLPAHLILGDMGGRFTNLXSLTVPFGKPNIDVTDAWDQ 300
QY 301 AMDAORIFKEAEKFFVSVGLPNNMTOGFWNSMLTDPGNVQKAVCHPAMDLGKDFRILM 360
D 301 AMDAORIFKEAEKFFVSVGLPNNMTOGFWNSMLTDPGNVQKAVCHPAMDLGKDFRILM 360
QY 361 CTKYTMDFLTAAHHEMGIQYDMAAOPFLLRNANGEGHEAVGEIMLSAATPKHLKS 420
D 361 CTKYTMDFLTAAHHEMGIQYDMAAOPFLLRNANGEGHEAVGEIMLSAATPKHLKS 420
QY 421 IGLSPDFOEDNETEINFLKQALTYGTLPFTYMLEKRWMPFKGEIPKQDMKKWEM 480
D 421 IGLSPDFOEDNETEINFLKQALTYGTLPFTYMLEKRWMPFKGEIPKQDMKKWEM 480
QY 481 KREIVGVPEVPDHECYCDPASLFHVSNDSFIRYTRTLXOFQOALCOAAKHGEPHL 540
D 481 KREIVGVPEVPDHECYCDPASLFHVSNDSFIRYTRTLXOFQOALCOAAKHGEPHL 540
QY 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVYGAKNMVRPLNFEPLFTWLKDQNK 600
D 541 KCDISNSTEAGOKLFNMLRLGKSEPTLALENVYGAKNMVRPLNFEPLFTWLKDQNK 600
QY 601 NSFVGMSTDMSPYADOSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVN 660
D 601 NSFVGMSTDMSPYADOSIKVRISLKSALGDKAYEMNDENMYLFRSSVAYAMROYFLKVN 660
QY 661 QMILFGEEDVRANLKRISFNFEVYAPKAVSDIIPTEVEKAIKRSRSRINDAFRLND 720
D 661 QMILFGEEDVRANLKRISFNFEVYAPKAVSDIIPTEVEKAIKRSRSRINDAFRLND 720
QY 721 SLEFLGIQPTLGPDPNPVSIMLIVGVVGVIVGIVILFTGIRDKKKKARSGENP 780
D 721 SLEFLGIQPTLGPDPNPVSIMLIVGVVGVIVGIVILFTGIRDKKKKARSGENP 780
QY 781 YASIDISKGNNPGFQNTDVOYTSF 805
D 781 YASIDISKGNNPGFQNTDVOYTSF 805

RESULT 3
US-09-969-384-13
; Sequence 13, Application US/09969384
; Publication No. US20020192749A1
; GENERAL INFORMATION:
; APPLICANT: Moore, et al.
; TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT055P1

```

```

: CURRENT APPLICATION NUMBER: US/09/969,384
: CURRENT FILING DATE: 2001-10-03
: PRIOR APPLICATION NUMBER: PCT/US01/10542
: PRIOR FILING DATE: 2001-04-02
: PRIOR APPLICATION NUMBER: 60/236,384
: PRIOR FILING DATE: 2000-09-29
: PRIOR APPLICATION NUMBER: 60/194,118
: PRIOR FILING DATE: 2000-04-03
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 13
: LENGTH: 711
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (219)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (240)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (499)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-969-384-13

```

```

Query Match      88.0%; Score 3775; DB 9; Length 711;
Best Local Similarity 99.3%; Pred. No. 1.8e-305;
Matches 704; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 62 MNNAGDKWSAFLEKQSTLAQWYPLQEIQLTVKQLQALQONGSSVLEDESKSRRLNTILN 121
DB 1 MNNAGDKWSAFLEKQSTLAQWYPLQEIQLTVKQLQALQONGSSVLEDESKSRRLNTILN 60
QY 122 TMSITSTGKVCNPDNPQECILLEPGLNEIMANSIDYNERLWAMESRSEVQKQRLPYE 181
DB 61 TMSITSTGKVCNPDNPQECILLEPGLNEIMANSIDYNERLWAMESRSEVQKQRLPYE 120
QY 182 EYVYLAKNEKARANHYEDGDYWRGDEYNGVDYDSRGQLEDEVEHTFEELIKLYEHLH 241
DB 121 EYVYLAKNEKARANHYEDGDYWRGDEYNGVDYDSRGQLEDEVEHTFEELIKLYEHLH 180
QY 242 AYVRKILMAAYPSYISPIGCLPAHLLDGMGRFWNTNLSLTVPFGQKRNIDVTAMDQ 301
DB 181 AYVRKILMAAYPSYISPIGCLPAHLLDGMGRFWNTNLSLTVPFGQKRNIDVTAMDQ 240
QY 302 WDAQRIFEAEKFEFVSGLPNTQGFWENSMILTPGNAVOKAVCHPTAMDLSKGFRLIMC 361
DB 241 WDAQRIFEAEKFEFVSGLPNTQGFWENSMILTPGNAVOKAVCHPTAMDLSKGFRLIMC 300
QY 362 TVVTMDDELTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMLSAATPKHLKSI 421
DB 301 TVVTMDDELTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMLSAATPKHLKSI 360
QY 422 GILSPDFQEDNETEINFLKQALITVGLPPTYMLEKRWAVFKGEIPKQMMKKWEMK 481
DB 361 GILSPDFQEDNETEINFLKQALITVGLPPTYMLEKRWAVFKGEIPKQMMKKWEMK 420
QY 482 RELVGVVEVPDHEITCDPASLFFHVSNDYSFIRYRTLYQFOFQALCQAAKHESGPLHK 541
DB 421 RELVGVVEVPDHEITCDPASLFFHVSNDYSFIRYRTLYQFOFQALCQAAKHESGPLHK 480
QY 542 CDSINSTEAGOKLFNMLRKGSSEPTLAEENVGAKNMVRLPLNFEELFTWLKQONKN 601
DB 481 CDSINSTEAGOKLFNMLRKGSSEPTLAEENVGAKNMVRLPLNFEELFTWLKQONKN 540
QY 602 SPVGMSTDMSPYADQSIKVRISLKSALGKAYEMWDEKYLFRSSVAIYAMROYETLKYKQ 661
DB 541 SPVGMSTDMSPYADQSIKVRISLKSALGKAYEMWDEKYLFRSSVAIYAMROYETLKYKQ 600
QY 662 MILFGEDVAVANIKRISFNFPVTAPKNSDIIPREVEKATIRMSRSRINDAFRLINDS 721
DB 601 MILFGEDVAVANIKRISFNFPVTAPKNSDIIPREVEKATIRMSRSRINDAFRLINDS 660

```

```

QY 722 LEFLGIQPTLGPNOPPVSIWLIVFGVWGVTVGIVILIFTGIRDRK 770
DB 661 LEFLGIQPTLGPNOPPVSIWLIVFGVWGVTVGIVILIFTGIRDRK 709

```

## RESULT 4

US-09-969-384-25

Sequence 25, Application US/09969384

Publication No. US20020192749A1

GENERAL INFORMATION:

APPLICANT: Moore, et al.

TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies

FILE REFERENCE: PTO55P1

CURRENT APPLICATION NUMBER: US/09/969,384

PRIOR APPLICATION NUMBER: PCT/US01/10542

PRIOR FILING DATE: 2001-10-03

PRIOR FILING DATE: 2001-04-02

PRIOR APPLICATION NUMBER: 60/236,384

PRIOR FILING DATE: 2000-09-29

PRIOR APPLICATION NUMBER: 60/194,118

PRIOR FILING DATE: 2000-04-03

NUMBER OF SEQ ID NOS: 27

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 25

LENGTH: 681

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (219)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (240)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (499)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-969-384-25

```

Query Match      84.9%; Score 3643; DB 9; Length 681;
Best Local Similarity 99.3%; Pred. No. 1.6e-294;
Matches 676; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 62 MNNAGDKWSAFLEKQSTLAQWYPLQEIQLTVKQLQALQONGSSVLEDESKSRRLNTILN 121
DB 1 MNNAGDKWSAFLEKQSTLAQWYPLQEIQLTVKQLQALQONGSSVLEDESKSRRLNTILN 60
QY 122 TMSITSTGKVCNPDNPQECILLEPGLNEIMANSIDYNERLWAMESRSEVQKQRLPYE 181
DB 61 TMSITSTGKVCNPDNPQECILLEPGLNEIMANSIDYNERLWAMESRSEVQKQRLPYE 120
QY 182 EYVYLAKNEKARANHYEDGDYWRGDEYNGVDYDSRGQLEDEVEHTFEELIKLYEHLH 241
DB 121 EYVYLAKNEKARANHYEDGDYWRGDEYNGVDYDSRGQLEDEVEHTFEELIKLYEHLH 180
QY 242 AYVRKILMAAYPSYISPIGCLPAHLLDGMGRFWNTNLSLTVPFGQKRNIDVTAMDQ 301
DB 181 AYVRKILMAAYPSYISPIGCLPAHLLDGMGRFWNTNLSLTVPFGQKRNIDVTAMDQ 240
QY 302 WDAQRIFEAEKFEFVSGLPNTQGFWENSMILTPGNAVOKAVCHPTAMDLSKGFRLIMC 361
DB 241 WDAQRIFEAEKFEFVSGLPNTQGFWENSMILTPGNAVOKAVCHPTAMDLSKGFRLIMC 300
QY 362 TVVTMDDELTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMLSAATPKHLKSI 421
DB 301 TVVTMDDELTAHHEMGHIQYDMAAOPFLLRNGANEGFHEAVGEIMLSAATPKHLKSI 360
QY 422 GILSPDFQEDNETEINFLKQALITVGLPPTYMLEKRWAVFKGEIPKQMMKKWEMK 481
DB 361 GILSPDFQEDNETEINFLKQALITVGLPPTYMLEKRWAVFKGEIPKQMMKKWEMK 420
QY 482 RELVGVVEVPDHEITCDPASLFFHVSNDYSFIRYRTLYQFOFQALCQAAKHESGPLHK 541

```

```

Db 421 REIVGVEPEVPHDETCDPASLFHVSNDYSFIRYTRTYXOFQFALCOAKHCEPLK 480
OY 542 CDISNSTEAGOKLFNNLRGKSEPMWLALENVGAKNMVRPLNFEPLFWLKDQKN 601
Db 481 CDISNSTEAGOKLFNNLRGKSEPMWLALENVGAKNMVRPLNFEPLFWLKDQKN 540
OY 602 SEYGMSTWSPYADOSIKVRISLKSALGDKAYEMDNEMYLFRRSSVAYAMROYFLKVKNO 661
Db 541 SEYGMSTWSPYADOSIKVRISLKSALGDKAYEMDNEMYLFRRSSVAYAMROYFLKVKNO 600
OY 662 MLFGEDEVANLKRISFNEFVYAPKAVSDIIPREVEKAIKMSRKRINDAFRLDNS 721
Db 601 MLFGEDEVANLKRISFNEFVYAPKAVSDIIPREVEKAIKMSRKRINDAFRLDNS 660
OY 722 LEFLGIQPTLGPPOPPVSIW 742
Db 661 LEFLGIQPTLGPPOPPVSIW 681

```

## RESULT 5

```

US-09-978-385-6
; Sequence 6, Application US/09978385
; Patent No. US20020177211A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Pectie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24C1
; CURRENT APPLICATION NUMBER: US/09/978,385
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/133,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Mouse
US-09-978-385-6

```

```

Query Match 83.4%; Score 3579; DB 9; Length 805;
Best Local Similarity 82.1%; Pred. No. 4.4e-289;
Matches 661; Conservative 60; Mismatches 84; Indels 0; Gaps 0;

```

```

OY 1 MSSSSWLLSLVAATAOSTIEQAKTFELDKFNHEADELFYQSSLASWNTNITEENVQ 60
Db 1 MSSSSWLLSLVAATAOSTIEQAKTFELDKFNHEADELFYQSSLASWNTNITEENVQ 60
OY 61 MNNAAGDKWSAFLEKQOSTLAQMPLOEIQNLTVKLOALQOQSSSVLSDEKSKRLNTIL 120
Db 61 KMSEAAKWSAFYEQSKTQOSFSLQEIPTPIIKROLQALQOQSSSVLSADKKKQJNTIL 120
OY 121 NTWSTIYSTGVCPNPNPOECLELLEPGLNEIMANSIDYNERLWAMWSRSEVQKOLRPLY 180
Db 121 NTWSTIYSTGVCPNPNPOECLELLEPGLNEIMANSIDYNERLWAMWSRSEVQKOLRPLY 180
OY 181 EEVYVLKNEARAHNYEDGDYRGDYEVANGVGYDYSKQGLIEDVHTPEETIKPLYEHL 240
Db 181 EEVYVLKNEARAHNYEDGDYRGDYEVANGVGYDYSKQGLIEDVHTPEETIKPLYEHL 240
OY 241 HAVYRAKLMAVPSYISPLIGCLPAHLGDMGHEFTNLSLVPEQCKINIVDTAMVQ 300
Db 241 HAVYRRKIMDTYPSYISPLIGCLPAHLGDMGHEFTNLSLVPEQCKINIVDTAMVQ 300
OY 301 AMDAQIIEAEKFFYVGLPNMTQGFWNSMLTDPGNVQKAVCHPTAWDLCKGDFRILM 360
Db 301 AMDAQIIEAEKFFYVGLPNMTQGFWNSMLTDPGNVQKAVCHPTAWDLCKGDFRILM 360

```

```

Db 301 GMDAERITQEAKEKFFYVGLPMTQGFWNSMLTEPADGRKVCYCHPTAWDLCKGDFRILM 360
OY 361 CTKYTMDEPLTAHHEMGHIQYDMAVAAOPFLLRNGANGCFHEAVGELTMSLSAATPKILKS 420
Db 361 CTKYTMDEPLTAHHEMGHIQYDMAVAAOPFLLRNGANGCFHEAVGELTMSLSAATPKILKS 420
OY 421 IGLSPQOEENETEFNLKQALITVGLTPEYTMLEKRMWYFKGIPRDKMKWEM 480
Db 421 IGLSPQOEENETEFNLKQALITVGLTPEYTMLEKRMWYFKGIPRDKMKWEM 480
OY 481 KREIVGVEPEVPHDETCDPASLFHVSNDYSFIRYTRTYXOFQFALCOAKHCEPLK 540
Db 481 KREIVGVEPEVPHDETCDPASLFHVSNDYSFIRYTRTYXOFQFALCOAKHCEPLK 540
OY 541 KCDISNSTEAGOKLFNNLRGKSEPMWLALENVGAKNMVRPLNFEPLFWLKDQKN 600
Db 541 KCDISNSTEAGOKLFNNLRGKSEPMWLALENVGAKNMVRPLNFEPLFWLKDQKN 600
OY 601 NSFVGMSTWSPYADOSIKVRISLKSALGDKAYEMDNEMYLFRRSSVAYAMROYFLKVKNO 660
Db 601 NSFVGMSTWSPYADOSIKVRISLKSALGDKAYEMDNEMYLFRRSSVAYAMROYFLKVKNO 660
OY 661 QMLFGEDEVANLKRISFNEFVYAPKAVSDIIPREVEKAIKMSRKRINDAFRLDNS 720
Db 661 QVPELEDEVYRSDLRKRVSFYFVTSPOVSDVIFRSEVEDAIRSKRINDAFRLDNS 720
OY 721 SLEFLGIQPTLGPPOPPVSIWLVFQVGVGIVGLIFGTIDRRKKKARSGENP 780
Db 721 SLEFLGIQPTLGPPOPPVSIWLVFQVGVGIVGLIFGTIDRRKKKARSGENP 780
OY 781 YASIDISGENNPGFQNTDVOYTSF 805
Db 781 YDSMDIGKGSNAGFQNSDDAQTSTF 805

```

## RESULT 6

```

US-09-978-385-9
; Sequence 9, Application US/09978385
; Patent No. US20020177211A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Pectie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24C1
; CURRENT APPLICATION NUMBER: US/09/978,385
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/133,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 805
; TYPE: PRT
; ORGANISM: Mouse
US-09-978-385-9

```

```

Query Match 83.0%; Score 3561; DB 9; Length 805;
Best Local Similarity 81.9%; Pred. No. 1.4e-287;
Matches 659; Conservative 60; Mismatches 86; Indels 0; Gaps 0;

```

```

OY 1 MSSSSWLLSLVAATAOSTIEQAKTFELDKFNHEADELFYQSSLASWNTNITEENVQ 60
Db 1 MSSSSWLLSLVAATAOSTIEQAKTFELDKFNHEADELFYQSSLASWNTNITEENVQ 60
OY 61 MNNAAGDKWSAFLEKQOSTLAQMPLOEIQNLTVKLOALQOQSSSVLSDEKSKRLNTIL 120
Db 61 KMSEAAKWSAFYEQSKTQOSFSLQEIPTPIIKROLQALQOQSSSVLSADKKKQJNTIL 120

```



QY 121 NTMSTIYSGKVCNPDNPOECILLEPGLNEIMANSIDYNERLWAMESWSEVKGOLRPLY 180  
121 NTMSTIYSGKVCNPDNPOECILLEPGLNEIMANSIDYNERLWAMESWSEVKGOLRPLY 180  
QY 181 EEEVVLKNEARANNHEDYGDYRGYEVNGVGYRSROLLLEDVHPEETKPLYEHL 240  
181 EEEVVLKNEARANNHEDYGDYRGYEVNGVGYRSROLLLEDVHPEETKPLYEHL 240  
QY 241 HAYVRAKIMANAPSYISIPICLPAHLILGDMWGRFWNLXSLTVPFQGNPIDVTDAMVDO 300  
241 HAYVRAKIMANAPSYISIPICLPAHLILGDMWGRFWNLXSLTVPFQGNPIDVTDAMVDO 300  
QY 301 AMDAQRIFKEAEKFFVYVGLPNNTOGWEMSLTDGNNVOKACHPLAWLGGDRPILM 360  
301 AMDAQRIFKEAEKFFVYVGLPNNTOGWEMSLTDGNNVOKACHPLAWLGGDRPILM 360  
QY 361 CTKTMDPFLTAHHEMHIQYDMAAQAQPLLRNGANECHFAVGEIMSLSATPKHLKS 420  
361 CTKTMDPFLTAHHEMHIQYDMAAQAQPLLRNGANECHFAVGEIMSLSATPKHLKS 420  
QY 421 IGLSPDEQEDNETEINFLKQALTYGLPFTYMLEKRWVFKGEIPKDOMKKWEM 480  
421 IGLSPDEQEDNETEINFLKQALTYGLPFTYMLEKRWVFKGEIPKDOMKKWEM 480  
QY 481 KREIVGYVEPYPHDETCDPASLFEHVSNDYSEFRTYRTLYQOPQALQAAKHSEPLH 540  
481 KREIVGYVEPYPHDETCDPASLFEHVSNDYSEFRTYRTLYQOPQALQAAKHSEPLH 540  
QY 541 KCDISNSTEAGOKLFNNMLRLGKSEPTLLEENYVGAKNMVRPLNFEPLFTWLKQONK 600  
541 KCDISNSTEAGOKLFNNMLRLGKSEPTLLEENYVGAKNMVRPLNFEPLFTWLKQONK 600  
QY 601 NSFGWNTWSPYADQSIKVRISLKSALGDKAYENDNEMYLFRSSVAYAMROYELKYN 660  
601 NSFGWNTWSPYADQSIKVRISLKSALGDKAYENDNEMYLFRSSVAYAMROYELKYN 660  
QY 661 QMIFGEEDVYANIKPRISNFEVTPAKNSDIIIPREVKALIRMSRSRINDPFLNDN 720  
661 QMIFGEEDVYANIKPRISNFEVTPAKNSDIIIPREVKALIRMSRSRINDPFLNDN 720  
QY 721 SLEFGIOPTLGPNOQPVSIWLVFGVGMVIVGIVILFTGIRPKKKKRSSENP 780  
721 SLEFGIOPTLGPNOQPVSIWLVFGVGMVIVGIVILFTGIRPKKKKRSSENP 780  
QY 781 YASIDISKGENNPFQONTDVQTSF 805  
781 YASIDISKGENNPFQONTDVQTSF 805

RESULT 7  
US-10-028-072-72  
Sequence 72, Application US/10028072  
Publication No. US20030004311A1  
GENERAL INFORMATION:  
APPLICANT: Baker Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tunas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang  
TITLE OF INVENTION:

FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/028, 072  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059836  
PRIOR FILING DATE: 1997-09-24  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062285  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062816  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063045  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063082  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/063127  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063327  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063329  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063550  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063561  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063704  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063733  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063735  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063738  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063755  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064248  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/064809  
PRIOR FILING DATE: 1997-11-07  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065846  
PRIOR FILING DATE: 1997-11-17  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066453  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066511  
PRIOR FILING DATE: 1997-11-24

```

PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026

```

```

PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

```

```

Query Match          69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 M S S S W L L S L V A V T A A O S T I E O A K T F L D K R N H A E D L F Y O S S L A S M N Y N T N T E E N V Q 60
Db 1 M S S S W L L S L V A V T A A O S T I E O A K T F L D K R N H A E D L F Y O S S L A S M N Y N T N T E E N V Q 60
QY 61 N N N A G D K W S A F L K R Q S T L A O M Y P L O E I O N L T V K L Q L A L O Q N S S V S E D K S R L N T I L 120
Db 61 N N N A G D K W S A F L K R Q S T L A O M Y P L O E I O N L T V K L Q L A L O Q N S S V S E D K S R L N T I L 120
QY 121 N M S T I Y S T G K V C N P D N Q E C L L E P G L N E T M A N S L D Y N E R L M A W S R S E V G K O L R P L Y 180
Db 121 N M S T I Y S T G K V C N P D N Q E C L L E P G L N E T M A N S L D Y N E R L M A W S R S E V G K O L R P L Y 180
QY 181 E E Y V U L K N E M A R A N H Y E D Y G D Y W R G D Y E V N G V D Y S R G O L I E D V E H T F E E I P L Y E H L 240
Db 181 E E Y V U L K N E M A R A N H Y E D Y G D Y W R G D Y E V N G V D Y S R G O L I E D V E H T F E E I P L Y E H L 240
QY 241 H A Y V R A K L M A N A P S Y I S P I G C L P A H L L G D M G R E W T N L Y S L T V P F G Q K P N I D V T D A M V D Q 300
Db 241 H A Y V R A K L M A N A P S Y I S P I G C L P A H L L G D M G R E W T N L Y S L T V P F G Q K P N I D V T D A M V D Q 300
QY 301 A M D A O R I F K E A K F P V S G L P M T O G F E W E N S M L D P G V O A V C H P T A M D I G K D F R I L M 360
Db 301 A M D A O R I F K E A K F P V S G L P M T O G F E W E N S M L D P G V O A V C H P T A M D I G K D F R I L M 360
QY 361 C T K V T M D F L T A H H E M G I O Y D M A Y A O P F L R N G A N G F E A V G E L M S I S A A T P K L K S 420
Db 361 C T K V T M D F L T A H H E M G I O Y D M A Y A O P F L R N G A N G F E A V G E L M S I S A A T P K L K S 420
QY 421 I G L S P D F O E D N E T E I N F L K O A L T I V G T P T Y M L E K W M V E K G E I P K D O W K K W E M 480
Db 421 I G L S P D F O E D N E T E I N F L K O A L T I V G T P T Y M L E K W M V E K G E I P K D O W K K W E M 480
QY 481 K R E I G V E P V P H D E T Y C D P A S L F H V S D Y S F I R Y T R T L Y O F O F O A L C O A K H E G P L H 540
Db 481 K R E I G V E P V P H D E T Y C D P A S L F H V S D Y S F I R Y T R T L Y O F O F O A L C O A K H E G P L H 540

```

OY 541 KCDISNSTEAGOKL 554  
Db 541 KCDISNSTEAGOKL 554

## RESULT 8

US-10-121-049-72  
; Sequence 72, Application US/10121049  
; Publication No. US20030022239A1  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C17  
; CURRENT APPLICATION NUMBER: US/10/121,049  
; CURRENT FILING DATE: 2002-04-12  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 72  
; LENGTH: 555  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-121-049-72

Query Match 69.4%; Score 2979; DB 5; Length 555;  
Best Local Similarity 99.8%; Pred. No. 1.9e-239;  
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSSSSMLLSLVAVTAOSTIEEOKTFLDKFNHEAEDLFYQSSSLASWNTNITEENVQ 60  
Db 1 MSSSSMLLSLVAVTAOSTIEEOKTFLDKFNHEAEDLFYQSSSLASWNTNITEENVQ 60  
OY 61 NNNNAGDKWSAFLEKQSTLAQMPLOEIONLVKIQLOALQONGSSVLSDESKRLNTIL 120  
Db 61 NNNNAGDKWSAFLEKQSTLAQMPLOEIONLVKIQLOALQONGSSVLSDESKRLNTIL 120  
OY 121 NTMSTIYTGKVCNPNPOBCLLEPGINEIMANSIDYNERLWAMESRSVEVKOLRPLY 180  
Db 121 NTMSTIYTGKVCNPNPOBCLLEPGINEIMANSIDYNERLWAMESRSVEVKOLRPLY 180  
OY 181 EBYVVLKNEMARANHEDYGYWRGDEYVNGVDYDYSRGQLEIDVEHTEFEIKPLYLEHL 240  
Db 181 EBYVVLKNEMARANHEDYGYWRGDEYVNGVDYDYSRGQLEIDVEHTEFEIKPLYLEHL 240  
OY 241 HAYYRAKLMNAYPSYISPGICLPAHLGDMWGRFTNLSLTVPPGOKPNIDVTDAVQ 300  
Db 241 HAYYRAKLMNAYPSYISPGICLPAHLGDMWGRFTNLSLTVPPGOKPNIDVTDAVQ 300  
OY 301 AMDAORIFKEAEKFFVSGLPNMTQGFWENSMITDPGNVOKAVCHPTAMDIGKGFRIIM 360  
Db 301 AMDAORIFKEAEKFFVSGLPNMTQGFWENSMITDPGNVOKAVCHPTAMDIGKGFRIIM 360  
OY 361 CTKYTMDDFLTAHHEMGIQYDAAVAAQPFLLRNGANEGFHEAVGELMISLAAATPKHLKS 420  
Db 361 CTKYTMDDFLTAHHEMGIQYDAAVAAQPFLLRNGANEGFHEAVGELMISLAAATPKHLKS 420

OY 421 IGLSPDFQEDNETEINFLKQALTIYGTLPFTYMLEKRWNVKGEIPKQDQMKRWEM 480  
Db 421 IGLSPDFQEDNETEINFLKQALTIYGTLPFTYMLEKRWNVKGEIPKQDQMKRWEM 480  
OY 481 KREIVGVEPVPHDETCDDPASLFHVSNDYSFIRYRTYTLVQFOFQALCOAAHNEGFLH 540  
Db 481 KREIVGVEPVPHDETCDDPASLFHVSNDYSFIRYRTYTLVQFOFQALCOAAHNEGFLH 540  
OY 541 KCDISNSTEAGOKL 554  
Db 541 KCDISNSTEAGOKL 554

## RESULT 9

US-10-123-904-72  
; Sequence 72, Application US/10123904  
; Publication No. US20030022328A1  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3330R1C54  
; CURRENT APPLICATION NUMBER: US/10/123,904  
; CURRENT FILING DATE: 2002-04-16  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 72  
; LENGTH: 555  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-123-904-72

Query Match 69.4%; Score 2979; DB 9; Length 555;  
Best Local Similarity 99.8%; Pred. No. 1.9e-239;  
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSSSSMLLSLVAVTAOSTIEEOKTFLDKFNHEAEDLFYQSSSLASWNTNITEENVQ 60  
Db 1 MSSSSMLLSLVAVTAOSTIEEOKTFLDKFNHEAEDLFYQSSSLASWNTNITEENVQ 60  
OY 61 NNNNAGDKWSAFLEKQSTLAQMPLOEIONLVKIQLOALQONGSSVLSDESKRLNTIL 120  
Db 61 NNNNAGDKWSAFLEKQSTLAQMPLOEIONLVKIQLOALQONGSSVLSDESKRLNTIL 120  
OY 121 NTMSTIYTGKVCNPNPOBCLLEPGINEIMANSIDYNERLWAMESRSVEVKOLRPLY 180  
Db 121 NTMSTIYTGKVCNPNPOBCLLEPGINEIMANSIDYNERLWAMESRSVEVKOLRPLY 180  
OY 181 EBYVVLKNEMARANHEDYGYWRGDEYVNGVDYDYSRGQLEIDVEHTEFEIKPLYLEHL 240  
Db 181 EBYVVLKNEMARANHEDYGYWRGDEYVNGVDYDYSRGQLEIDVEHTEFEIKPLYLEHL 240  
OY 241 HAYYRAKLMNAYPSYISPGICLPAHLGDMWGRFTNLSLTVPPGOKPNIDVTDAVQ 300  
Db 241 HAYYRAKLMNAYPSYISPGICLPAHLGDMWGRFTNLSLTVPPGOKPNIDVTDAVQ 300  
OY 301 AMDAORIFKEAEKFFVSGLPNMTQGFWENSMITDPGNVOKAVCHPTAMDIGKGFRIIM 360

```
|||||
Db 301 AMDAQRIFKEAEKFFVSGVGLPNTQGFWEKSMITDPGNVOKAVCHPTAMDLCRGDFRILM 360
Qy 361 CTKYVMDDEFLTAHHEMGHIQYDMAVYAAOPFLLRNGANEGFHEAVGELMSAATPKHLKS 420
Db 361 CTKYVMDDEFLTAHHEMGHIQYDMAVYAAOPFLLRNGANEGFHEAVGELMSAATPKHLKS 420
Qy 421 IGLSPDPOEDNETEINFLKQALITVGTLPFTYMLEKRWMMVFKEGELPKDQMKKKWEM 480
Db 421 IGLSPDPOEDNETEINFLKQALITVGTLPFTYMLEKRWMMVFKEGELPKDQMKKKWEM 480
Qy 481 KREIVGVEPVPHDETCDPASLFHVSNDYSFIRYTRTLXQFOFOALCOAAKHGEPIL 540
Db 481 KREIVGVEPVPHDETCDPASLFHVSNDYSFIRYTRTLXQFOFOALCOAAKHGEPIL 540
Qy 541 KCDISNSTEAGOKL 554
Db 541 KCDISNSTEAGOKL 554
```

## RESULT 10

```
US-10-140-470-72
; Sequence 72, Application US/10140470
; Publication No. US2003002231A1
```

## GENERAL INFORMATION:

```
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; PRIOR FILING DATE: 2002-05-06
; PRIOR APPLICATION removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-470-72
```

## Query Match

69.4%; Score 2979; DB 9; Length 555;

Best Local Similarity 99.8%; Pred. No. 1, 9e-239; Mismatches 0; Indels 0; Gaps 0;

```
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSSSSWLLSLVAVTAASSTIEBOAKTFLDKFNEHAEDELFYOSSLASMNNTNTEENVQ 60
Db 1 MSSSSWLLSLVAVTAASSTIEBOAKTFLDKFNEHAEDELFYOSSLASMNNTNTEENVQ 60
Qy 61 MNMNAAGKWSAFLKEOSTLAQMPLOEIONLTVKLOLQALQONSSSVLSSEDKSKRLNTIL 120
Db 61 MNMNAAGKWSAFLKEOSTLAQMPLOEIONLTVKLOLQALQONSSSVLSSEDKSKRLNTIL 120
Qy 121 NTMSTIYSTGKVCNPDPQECILLEPGLNEIMANSIDYNERLWAMSWRSEVGKOLRPLY 180
Db 121 NTMSTIYSTGKVCNPDPQECILLEPGLNEIMANSIDYNERLWAMSWRSEVGKOLRPLY 180
Qy 181 EETVVLAKENARANHEDYDGYRGDYEVNGVGYTYSKGLTEYVHFFEEKPLYEHL 240
Db 181 EETVVLAKENARANHEDYDGYRGDYEVNGVGYTYSKGLTEYVHFFEEKPLYEHL 240
```

```
Db 181 EETVVLAKENARANHEDYDGYRGDYEVNGVGYTYSKGLTEYVHFFEEKPLYEHL 240
Qy 241 HAYVRAKIMAYSYSPIGCLPAHLIGMKGFEWNTLSLVYRPGQKPNIDVTAMDQ 300
Db 241 HAYVRAKIMAYSYSPIGCLPAHLIGMKGFEWNTLSLVYRPGQKPNIDVTAMDQ 300
Qy 301 AMDAQRIFKEAEKFFVSGVGLPNTQGFWEKSMITDPGNVOKAVCHPTAMDLCRGDFRILM 360
Db 301 AMDAQRIFKEAEKFFVSGVGLPNTQGFWEKSMITDPGNVOKAVCHPTAMDLCRGDFRILM 360
Qy 361 CTKYVMDDEFLTAHHEMGHIQYDMAVYAAOPFLLRNGANEGFHEAVGELMSAATPKHLKS 420
Db 361 CTKYVMDDEFLTAHHEMGHIQYDMAVYAAOPFLLRNGANEGFHEAVGELMSAATPKHLKS 420
Qy 421 IGLSPDPOEDNETEINFLKQALITVGTLPFTYMLEKRWMMVFKEGELPKDQMKKKWEM 480
Db 421 IGLSPDPOEDNETEINFLKQALITVGTLPFTYMLEKRWMMVFKEGELPKDQMKKKWEM 480
Qy 481 KREIVGVEPVPHDETCDPASLFHVSNDYSFIRYTRTLXQFOFOALCOAAKHGEPIL 540
Db 481 KREIVGVEPVPHDETCDPASLFHVSNDYSFIRYTRTLXQFOFOALCOAAKHGEPIL 540
Qy 541 KCDISNSTEAGOKL 554
Db 541 KCDISNSTEAGOKL 554
```

## RESULT 11

```
US-10-175-746-72
; Sequence 72, Application US/10175746
; Publication No. US20030027270A1
```

## GENERAL INFORMATION:

```
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-746-72
```

## Query Match

69.4%; Score 2979; DB 9; Length 555;

Best Local Similarity 99.8%; Pred. No. 1, 9e-239; Mismatches 0; Indels 0; Gaps 0;

```
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSSSSWLLSLVAVTAASSTIEBOAKTFLDKFNEHAEDELFYOSSLASMNNTNTEENVQ 60
Db 1 MSSSSWLLSLVAVTAASSTIEBOAKTFLDKFNEHAEDELFYOSSLASMNNTNTEENVQ 60
Qy 61 MNMNAAGKWSAFLKEOSTLAQMPLOEIONLTVKLOLQALQONSSSVLSSEDKSKRLNTIL 120
Db 61 MNMNAAGKWSAFLKEOSTLAQMPLOEIONLTVKLOLQALQONSSSVLSSEDKSKRLNTIL 120
```

```

QY 121 NTMSTIYTGKVCNPDNPOECLELLEPGLINEIMANSLDYNERLWAMESRSEVGKOLRPLY 180
DB 121 NMSTIYTGKVCNPDNPOECLELLEPGLINEIMANSLDYNERLWAMESRSEVGKOLRPLY 180
QY 181 EBYVVLKNEMARANHYEDYGDYWRGDEYVNGVDGYDYSRGQLIEDVEHTFEELKPLYEHL 240
DB 181 EBYVVLKNEMARANHYEDYGDYWRGDEYVNGVDGYDYSRGQLIEDVEHTFEELKPLYEHL 240
QY 241 HAYVRAKIMANAPSYISIPICLPAHLIGDMWGRFTNLISLTPFGOKPNIDVTAMDQ 300
DB 241 HAYVRAKIMANAPSYISIPICLPAHLIGDMWGRFTNLISLTPFGOKPNIDVTAMDQ 300
QY 301 AMDAORIFKEAEKFEFVSGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGKDFRILM 360
DB 301 AMDAORIFKEAEKFEFVSGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGKDFRILM 360
QY 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAPPELLRNGANEGHEAVGEIMSLSAATPKHLKS 420
DB 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAPPELLRNGANEGHEAVGEIMSLSAATPKHLKS 420
QY 421 IGLLSPDQEDNTEINFLKQALITVGLTPTMYLKKRMWYFKGEIPKQDMKKWEM 480
DB 421 IGLLSPDQEDNTEINFLKQALITVGLTPTMYLKKRMWYFKGEIPKQDMKKWEM 480
QY 481 KREIYGVEPVPHDETYCPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAKHEGPLH 540
DB 481 KREIYGVEPVPHDETYCPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAKHEGPLH 540
QY 541 KCDISNSTEAGOKL 554
DB 541 KCDISNSTEAGOKL 554

```

## RESULT 12

US-10-176-918-72

Sequence 72, Application US/10176918

Publication No. US20030027275A1

GENERAL INFORMATION:

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Deforge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P330R1C382
: CURRENT APPLICATION NUMBER: US/10/176,918
: PRIORITY FILING DATE: 2002-06-20
: Prior Application removed - See File Wrapper or Palm
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 72
: LENGTH: 555
: TYPE: PRT
: ORGANISM: Homo Sapien
US-10-176-918-72

```

Query Match 69.4%; Score 2979; DB 9; Length 555;  
 Best Local Similarity 99.8%; Pred. No. 1.9e-239;  
 Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSSSSWLLSLVAVTAAQSTIEQAKTFLLKFNHEADLFYQSSLASNNTNITEENQ 60
DB 1 MSSSSWLLSLVAVTAAQSTIEQAKTFLLKFNHEADLFYQSSLASNNTNITEENQ 60
QY 61 MNMAGDKWSAFLEKQSTLAQNTPLQEIOMLYKLOLQALQNGSSVLSDEKSKRINTLI 120
DB 61 MNMAGDKWSAFLEKQSTLAQNTPLQEIOMLYKLOLQALQNGSSVLSDEKSKRINTLI 120
QY 121 NTMSTIYTGKVCNPDNPOECLELLEPGLINEIMANSLDYNERLWAMESRSEVGKOLRPLY 180
DB 121 NTMSTIYTGKVCNPDNPOECLELLEPGLINEIMANSLDYNERLWAMESRSEVGKOLRPLY 180
QY 181 EBYVVLKNEMARANHYEDYGDYWRGDEYVNGVDGYDYSRGQLIEDVEHTFEELKPLYEHL 240
DB 181 EBYVVLKNEMARANHYEDYGDYWRGDEYVNGVDGYDYSRGQLIEDVEHTFEELKPLYEHL 240
QY 241 HAYVRAKIMANAPSYISIPICLPAHLIGDMWGRFTNLISLTPFGOKPNIDVTAMDQ 300
DB 241 HAYVRAKIMANAPSYISIPICLPAHLIGDMWGRFTNLISLTPFGOKPNIDVTAMDQ 300
QY 301 AMDAORIFKEAEKFEFVSGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGKDFRILM 360
DB 301 AMDAORIFKEAEKFEFVSGLPNMTQGFWENSMITDPGNVQKAVCHPTAMDLGKDFRILM 360
QY 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAPPELLRNGANEGHEAVGEIMSLSAATPKHLKS 420
DB 361 CTKYTMDDFLTAHHEMGHIQYDMAVAAPPELLRNGANEGHEAVGEIMSLSAATPKHLKS 420
QY 421 IGLLSPDQEDNTEINFLKQALITVGLTPTMYLKKRMWYFKGEIPKQDMKKWEM 480
DB 421 IGLLSPDQEDNTEINFLKQALITVGLTPTMYLKKRMWYFKGEIPKQDMKKWEM 480
QY 481 KREIYGVEPVPHDETYCPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAKHEGPLH 540
DB 481 KREIYGVEPVPHDETYCPASLFHVSNDYSFIIRYTRTLXQFOFQALCOAKHEGPLH 540
QY 541 KCDISNSTEAGOKL 554
DB 541 KCDISNSTEAGOKL 554

```

## RESULT 13

US-10-176-921-72

Sequence 72, Application US/10176921

Publication No. US20030027276A1

GENERAL INFORMATION:

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Deforge, Laura
: APPLICANT: Desnoyers, Luc
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P330R1C288
: CURRENT APPLICATION NUMBER: US/10/176,921
: PRIORITY FILING DATE: 2002-06-20
: Prior Application removed - See File Wrapper or Palm
: SEQ ID NO 72
: LENGTH: 555
: TYPE: PRT

```

ORGANISM: Homo Sapien  
us-10-176-921-72

Query Match 69.4%; Score 2979; DB 9; Length 555;  
Best Local Similarity 99.8%; Pred. No. 1.9e-239;  
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSSSSWLLSLVAVTAAGSTIEBOAKTFLEKFNHEADLFYQSSLASMNNTNTEENVQ 60
   1 MSSSSWLLSLVAVTAAGSTIEBOAKTFLEKFNHEADLFYQSSLASMNNTNTEENVQ 60
Db 1 MSSSSWLLSLVAVTAAGSTIEBOAKTFLEKFNHEADLFYQSSLASMNNTNTEENVQ 60
QY 61 MNMAGDKWSAFLEKQSTLAQMPLOEIQLTVKLOLQALQNGSSVLSDEKSKRLNTIL 120
   61 MNMAGDKWSAFLEKQSTLAQMPLOEIQLTVKLOLQALQNGSSVLSDEKSKRLNTIL 120
Db 61 MNMAGDKWSAFLEKQSTLAQMPLOEIQLTVKLOLQALQNGSSVLSDEKSKRLNTIL 120
QY 121 NTMSTIYTGKVCNPDNPQECLEPGLNEIMANSIDYNERLWAMESRSEVKGQRLPLY 180
   121 NTMSTIYTGKVCNPDNPQECLEPGLNEIMANSIDYNERLWAMESRSEVKGQRLPLY 180
Db 121 NTMSTIYTGKVCNPDNPQECLEPGLNEIMANSIDYNERLWAMESRSEVKGQRLPLY 180
QY 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRGQLIEDYHTEFEIKPLYEHL 240
   181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRGQLIEDYHTEFEIKPLYEHL 240
Db 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRGQLIEDYHTEFEIKPLYEHL 240
QY 241 HAYVRAKLMNAYPSYISPIGCLPAHLGDMGREFWNTLYSLTVPGQKPNIDVTDAVQ 300
   241 HAYVRAKLMNAYPSYISPIGCLPAHLGDMGREFWNTLYSLTVPGQKPNIDVTDAVQ 300
Db 241 HAYVRAKLMNAYPSYISPIGCLPAHLGDMGREFWNTLYSLTVPGQKPNIDVTDAVQ 300
QY 301 AMDAQRIFKEAEKFFVSVGLPMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRIIM 360
   301 AMDAQRIFKEAEKFFVSVGLPMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRIIM 360
Db 301 AMDAQRIFKEAEKFFVSVGLPMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRIIM 360
QY 361 CTKVTMDDEFLTAHHEMGIQYDMAYAAOPFLRNGANECEGHEAVGEIMSLAATPKHLKS 420
   361 CTKVTMDDEFLTAHHEMGIQYDMAYAAOPFLRNGANECEGHEAVGEIMSLAATPKHLKS 420
Db 361 CTKVTMDDEFLTAHHEMGIQYDMAYAAOPFLRNGANECEGHEAVGEIMSLAATPKHLKS 420
QY 421 IGLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWMPFKGEIIPKQDMKKWEM 480
   421 IGLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWMPFKGEIIPKQDMKKWEM 480
Db 421 IGLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWMPFKGEIIPKQDMKKWEM 480
QY 481 KREIVGVEPVPHDETCDPASLFHVSNDSFIRYTRTLVYQFOFQALCOAAKHGEPHL 540
   481 KREIVGVEPVPHDETCDPASLFHVSNDSFIRYTRTLVYQFOFQALCOAAKHGEPHL 540
Db 481 KREIVGVEPVPHDETCDPASLFHVSNDSFIRYTRTLVYQFOFQALCOAAKHGEPHL 540
QY 541 KCDISNSTEAGOKL 554
   541 KCDISNSTEAGOKL 554
Db 541 KCDISNSTEAGOKL 554

```

RESULT 14  
US-10-137-865-72  
; Sequence 72, Application US/10137865  
; Publication No. US20030032155A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C154

CURRENT APPLICATION NUMBER: US/10/137,865  
; CURRENT FILING DATE: 2002-05-03  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 72  
; LENGTH: 555  
; TYPE: PRF  
; ORGANISM: Homo Sapien  
us-10-137-865-72

Query Match 69.4%; Score 2979; DB 9; Length 555;  
Best Local Similarity 99.8%; Pred. No. 1.9e-239;  
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSSSSWLLSLVAVTAAGSTIEBOAKTFLEKFNHEADLFYQSSLASMNNTNTEENVQ 60
   1 MSSSSWLLSLVAVTAAGSTIEBOAKTFLEKFNHEADLFYQSSLASMNNTNTEENVQ 60
Db 1 MSSSSWLLSLVAVTAAGSTIEBOAKTFLEKFNHEADLFYQSSLASMNNTNTEENVQ 60
QY 61 MNMAGDKWSAFLEKQSTLAQMPLOEIQLTVKLOLQALQNGSSVLSDEKSKRLNTIL 120
   61 MNMAGDKWSAFLEKQSTLAQMPLOEIQLTVKLOLQALQNGSSVLSDEKSKRLNTIL 120
Db 61 MNMAGDKWSAFLEKQSTLAQMPLOEIQLTVKLOLQALQNGSSVLSDEKSKRLNTIL 120
QY 121 NTMSTIYTGKVCNPDNPQECLEPGLNEIMANSIDYNERLWAMESRSEVKGQRLPLY 180
   121 NTMSTIYTGKVCNPDNPQECLEPGLNEIMANSIDYNERLWAMESRSEVKGQRLPLY 180
Db 121 NTMSTIYTGKVCNPDNPQECLEPGLNEIMANSIDYNERLWAMESRSEVKGQRLPLY 180
QY 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRGQLIEDYHTEFEIKPLYEHL 240
   181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRGQLIEDYHTEFEIKPLYEHL 240
Db 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRGQLIEDYHTEFEIKPLYEHL 240
QY 241 HAYVRAKLMNAYPSYISPIGCLPAHLGDMGREFWNTLYSLTVPGQKPNIDVTDAVQ 300
   241 HAYVRAKLMNAYPSYISPIGCLPAHLGDMGREFWNTLYSLTVPGQKPNIDVTDAVQ 300
Db 241 HAYVRAKLMNAYPSYISPIGCLPAHLGDMGREFWNTLYSLTVPGQKPNIDVTDAVQ 300
QY 301 AMDAQRIFKEAEKFFVSVGLPMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRIIM 360
   301 AMDAQRIFKEAEKFFVSVGLPMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRIIM 360
Db 301 AMDAQRIFKEAEKFFVSVGLPMTQGFWENSMITDPGNVQKAVCHPTAMDLGKGFRIIM 360
QY 361 CTKVTMDDEFLTAHHEMGIQYDMAYAAOPFLRNGANECEGHEAVGEIMSLAATPKHLKS 420
   361 CTKVTMDDEFLTAHHEMGIQYDMAYAAOPFLRNGANECEGHEAVGEIMSLAATPKHLKS 420
Db 361 CTKVTMDDEFLTAHHEMGIQYDMAYAAOPFLRNGANECEGHEAVGEIMSLAATPKHLKS 420
QY 421 IGLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWMPFKGEIIPKQDMKKWEM 480
   421 IGLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWMPFKGEIIPKQDMKKWEM 480
Db 421 IGLSPDFQEDNETEINFLKQALITVGLPFTYMLEKRWMPFKGEIIPKQDMKKWEM 480
QY 481 KREIVGVEPVPHDETCDPASLFHVSNDSFIRYTRTLVYQFOFQALCOAAKHGEPHL 540
   481 KREIVGVEPVPHDETCDPASLFHVSNDSFIRYTRTLVYQFOFQALCOAAKHGEPHL 540
Db 481 KREIVGVEPVPHDETCDPASLFHVSNDSFIRYTRTLVYQFOFQALCOAAKHGEPHL 540
QY 541 KCDISNSTEAGOKL 554
   541 KCDISNSTEAGOKL 554
Db 541 KCDISNSTEAGOKL 554

```

RESULT 15  
US-10-140-474-72  
; Sequence 72, Application US/10140474  
; Publication No. US20030032156A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.

```

; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C162
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-474-72

Query Match          69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSLVAVTAOSTIEBOAKTFLDKFNHEAEDLFYQSSLASMYNTNITEENVQ 60
Db 1 MSSSSWLLSLVAVTAOSTIEBOAKTFLDKFNHEAEDLFYQSSLASMYNTNITEENVQ 60
QY 61 MNNAAGDKWSAFLEKQSTLAQMPLOEIONLVKLOLQALQONGSSVLSSEKSKRLNTIL 120
Db 61 MNNAAGDKWSAFLEKQSTLAQMPLOEIONLVKLOLQALQONGSSVLSSEKSKRLNTIL 120
QY 121 NTMSTIYSGKVCNPNPQECILLEPGLNEIMANSIDYNERLMAWMSRSEVQKOLRPLY 180
Db 121 NTMSTIYSGKVCNPNPQECILLEPGLNEIMANSIDYNERLMAWMSRSEVQKOLRPLY 180
QY 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRQGLIEDVEHTFEETIKPLYEHL 240
Db 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRQGLIEDVEHTFEETIKPLYEHL 240
QY 241 HAYYRAKIMNAYPSYISPIGCLPAHLGDMWGRFWTNLYSLVPPGOKPNIDVTDAVQ 300
Db 241 HAYYRAKIMNAYPSYISPIGCLPAHLGDMWGRFWTNLYSLVPPGOKPNIDVTDAVQ 300
QY 301 AMDAQRIFKEAEKEFFVSGLPNMTQGFWNSMLTDPGNVQKAVCHPTAMDLGKGFRIIL 360
Db 301 AMDAQRIFKEAEKEFFVSGLPNMTQGFWNSMLTDPGNVQKAVCHPTAMDLGKGFRIIL 360
QY 361 CTKYTMDDFLTAHHEMGIQYDMAAOPFLLRNGANEGFHAAGEIMSLSAATPKHLKS 420
Db 361 CTKYTMDDFLTAHHEMGIQYDMAAOPFLLRNGANEGFHAAGEIMSLSAATPKHLKS 420
QY 421 IGLSPFOEDNETEINFLLKQALTYGTLPTTYMLEKMRMVFKEGIPKQOMKKWEM 480
Db 421 IGLSPFOEDNETEINFLLKQALTYGTLPTTYMLEKMRMVFKEGIPKQOMKKWEM 480
QY 481 KREIVGVPEVPDHETTCBPASLFHVSNDYSFIRYTRTLVYQFOEALCOAKHHEGPLH 540
Db 481 KREIVGVPEVPDHETTCBPASLFHVSNDYSFIRYTRTLVYQFOEALCOAKHHEGPLH 540
QY 541 KCDISNSTEAGQKL 554
Db 541 KCDISNSTEAGQKL 554

RESULT 16
US-10-142-431-72
; Sequence 72, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang

```

```

; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C251
; CURRENT FILING DATE: 2002-05-10
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 72
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-142-431-72

Query Match          69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSLVAVTAOSTIEBOAKTFLDKFNHEAEDLFYQSSLASMYNTNITEENVQ 60
Db 1 MSSSSWLLSLVAVTAOSTIEBOAKTFLDKFNHEAEDLFYQSSLASMYNTNITEENVQ 60
QY 61 MNNAAGDKWSAFLEKQSTLAQMPLOEIONLVKLOLQALQONGSSVLSSEKSKRLNTIL 120
Db 61 MNNAAGDKWSAFLEKQSTLAQMPLOEIONLVKLOLQALQONGSSVLSSEKSKRLNTIL 120
QY 121 NTMSTIYSGKVCNPNPQECILLEPGLNEIMANSIDYNERLMAWMSRSEVQKOLRPLY 180
Db 121 NTMSTIYSGKVCNPNPQECILLEPGLNEIMANSIDYNERLMAWMSRSEVQKOLRPLY 180
QY 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRQGLIEDVEHTFEETIKPLYEHL 240
Db 181 EEEVVLKNEKMARANHEDYGDYWRGDEYVNGVDGYDSRQGLIEDVEHTFEETIKPLYEHL 240
QY 241 HAYYRAKIMNAYPSYISPIGCLPAHLGDMWGRFWTNLYSLVPPGOKPNIDVTDAVQ 300
Db 241 HAYYRAKIMNAYPSYISPIGCLPAHLGDMWGRFWTNLYSLVPPGOKPNIDVTDAVQ 300
QY 301 AMDAQRIFKEAEKEFFVSGLPNMTQGFWNSMLTDPGNVQKAVCHPTAMDLGKGFRIIL 360
Db 301 AMDAQRIFKEAEKEFFVSGLPNMTQGFWNSMLTDPGNVQKAVCHPTAMDLGKGFRIIL 360
QY 361 CTKYTMDDFLTAHHEMGIQYDMAAOPFLLRNGANEGFHAAGEIMSLSAATPKHLKS 420
Db 361 CTKYTMDDFLTAHHEMGIQYDMAAOPFLLRNGANEGFHAAGEIMSLSAATPKHLKS 420
QY 421 IGLSPFOEDNETEINFLLKQALTYGTLPTTYMLEKMRMVFKEGIPKQOMKKWEM 480
Db 421 IGLSPFOEDNETEINFLLKQALTYGTLPTTYMLEKMRMVFKEGIPKQOMKKWEM 480
QY 481 KREIVGVPEVPDHETTCBPASLFHVSNDYSFIRYTRTLVYQFOEALCOAKHHEGPLH 540
Db 481 KREIVGVPEVPDHETTCBPASLFHVSNDYSFIRYTRTLVYQFOEALCOAKHHEGPLH 540
QY 541 KCDISNSTEAGQKL 554
Db 541 KCDISNSTEAGQKL 554

RESULT 17
US-10-143-114-72
; Sequence 72, Application US/10143114
; Publication No. US20030036180A1

```



```

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C211
CURRENT APPLICATION NUMBER: US/10/143,114
CURRENT FILING DATE: 2002-05-09
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-143-114-72

```

```

Query Match      69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSSSMLLSLVAVTAOSTIEEQAFTLDFKNHEADLFYQSSLSAWNTNTTEENVQ 60
DB 1 MSSSSMLLSLVAVTAOSTIEEQAFTLDFKNHEADLFYQSSLSAWNTNTTEENVQ 60
QY 61 MNMNAAGKMSAFLEQSTLAQMPPLQEIQLNLTVKQLQALQNGSSVLSDEKSKRLNTIL 120
DB 61 MNMNAAGKMSAFLEQSTLAQMPPLQEIQLNLTVKQLQALQNGSSVLSDEKSKRLNTIL 120
QY 121 NTMSTIYSTGKVCNPDPQECLEPGLNEIMANSLDYNERLMAWMSRSEVKGKRLPLY 180
DB 121 NTMSTIYSTGKVCNPDPQECLEPGLNEIMANSLDYNERLMAWMSRSEVKGKRLPLY 180
QY 181 EEEYVLKNEAMARAHNYEDYGDYWRGDIYVNGVDYDYSRQGLLEDEYHTFEETIKPLYLEHL 240
DB 181 EEEYVLKNEAMARAHNYEDYGDYWRGDIYVNGVDYDYSRQGLLEDEYHTFEETIKPLYLEHL 240
QY 241 HAYYRAKLMAAHPYSYISPIGCLPAHLIGDMGREFWNTLSLVTFPGOKPNIDVTAMVQ 300
DB 241 HAYYRAKLMAAHPYSYISPIGCLPAHLIGDMGREFWNTLSLVTFPGOKPNIDVTAMVQ 300
QY 301 AMDAQRIFKEAEKFFVSVGLPNNTOGFWMENSLTDPGNVOKAVCHPTAMDLGKGRFRLIM 360
DB 301 AMDAQRIFKEAEKFFVSVGLPNNTOGFWMENSLTDPGNVOKAVCHPTAMDLGKGRFRLIM 360
QY 361 CTKYTMDDFLTAHHEMGHIQYDMAAOPFLRLNGANEGHEAVGEIMSLSATPRHLKS 420
DB 361 CTKYTMDDFLTAHHEMGHIQYDMAAOPFLRLNGANEGHEAVGEIMSLSATPRHLKS 420
QY 421 IGLSPDFQEDNTEINFLLKQALITVGLPTFTYMLEKRWMMVFGEIPKQDMKKWEM 480
DB 421 IGLSPDFQEDNTEINFLLKQALITVGLPTFTYMLEKRWMMVFGEIPKQDMKKWEM 480
QY 481 KREIVGVEPVPHEDETCYPASLPHVSNDSYFIRYTRTLVYQFOEALCQAAKHGGPLH 540
DB 481 KREIVGVEPVPHEDETCYPASLPHVSNDSYFIRYTRTLVYQFOEALCQAAKHGGPLH 540
QY 541 KCDISNSTAGOKL 554
DB 541 KCDISNSTAGOKL 554

```

DB 541 KCDISNSTAGOKL 554

```

RESULT 18
US-10-140-002-72
Sequence 72, Application US/10140002
Publication No. US20030037623A1
GENERAL INFORMATION:

```

```

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C059
CURRENT APPLICATION NUMBER: US/10/140,002
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 72
LENGTH: 555
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-002-72

```

```

Query Match      69.4%; Score 2979; DB 9; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.9e-239;
Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MSSSSMLLSLVAVTAOSTIEEQAFTLDFKNHEADLFYQSSLSAWNTNTTEENVQ 60
DB 1 MSSSSMLLSLVAVTAOSTIEEQAFTLDFKNHEADLFYQSSLSAWNTNTTEENVQ 60
QY 61 MNMNAAGKMSAFLEQSTLAQMPPLQEIQLNLTVKQLQALQNGSSVLSDEKSKRLNTIL 120
DB 61 MNMNAAGKMSAFLEQSTLAQMPPLQEIQLNLTVKQLQALQNGSSVLSDEKSKRLNTIL 120
QY 121 NTMSTIYSTGKVCNPDPQECLEPGLNEIMANSLDYNERLMAWMSRSEVKGKRLPLY 180
DB 121 NTMSTIYSTGKVCNPDPQECLEPGLNEIMANSLDYNERLMAWMSRSEVKGKRLPLY 180
QY 181 EEEYVLKNEAMARAHNYEDYGDYWRGDIYVNGVDYDYSRQGLLEDEYHTFEETIKPLYLEHL 240
DB 181 EEEYVLKNEAMARAHNYEDYGDYWRGDIYVNGVDYDYSRQGLLEDEYHTFEETIKPLYLEHL 240
QY 241 HAYYRAKLMAAHPYSYISPIGCLPAHLIGDMGREFWNTLSLVTFPGOKPNIDVTAMVQ 300
DB 241 HAYYRAKLMAAHPYSYISPIGCLPAHLIGDMGREFWNTLSLVTFPGOKPNIDVTAMVQ 300
QY 301 AMDAQRIFKEAEKFFVSVGLPNNTOGFWMENSLTDPGNVOKAVCHPTAMDLGKGRFRLIM 360
DB 301 AMDAQRIFKEAEKFFVSVGLPNNTOGFWMENSLTDPGNVOKAVCHPTAMDLGKGRFRLIM 360
QY 361 CTKYTMDDFLTAHHEMGHIQYDMAAOPFLRLNGANEGHEAVGEIMSLSATPRHLKS 420
DB 361 CTKYTMDDFLTAHHEMGHIQYDMAAOPFLRLNGANEGHEAVGEIMSLSATPRHLKS 420
QY 421 IGLSPDFQEDNTEINFLLKQALITVGLPTFTYMLEKRWMMVFGEIPKQDMKKWEM 480
DB 421 IGLSPDFQEDNTEINFLLKQALITVGLPTFTYMLEKRWMMVFGEIPKQDMKKWEM 480

```

QY 481 KREIVGVPEVPHDETCDBPASFHVNSDYSFIRYTRTLXQFOFOALCOAAKHGEPH 540  
 DB 481 KREIVGVPEVPHDETCDBPASFHVNSDYSFIRYTRTLXQFOFOALCOAAKHGEPH 540  
 QY 541 KCDISNSTEAGOKL 554  
 DB 541 KCDISNSTEAGOKL 554

## RESULT 19

US-10-142-419-72  
 ; Sequence 72, Application US/10142419  
 ; Publication No. US20030044945A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: Deforge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P3330R1C244  
 ; CURRENT APPLICATION NUMBER: US/10/142,419  
 ; CURRENT FILING DATE: 2002-05-10  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 72  
 ; LENGTH: 555  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien  
 US-10-142-419-72

Query Match 69.4%; Score 2979; DB %: Length 555;  
 Best Local Similarity 99.8%; Pred. No. 1.9e-239;  
 Matches 553; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSSWLLSLVAVTAOSTIEEQATFLDFKNHEADLFYQSSLASWNTNTITEENVQ 60  
 DB 1 MSSSSWLLSLVAVTAOSTIEEQATFLDFKNHEADLFYQSSLASWNTNTITEENVQ 60  
 QY 61 MNNAAGDKWSAFLEKQSTLAQMPLOETQNTLVKLOLQALQONSSVLSSEKSKRLNTIL 120  
 DB 61 MNNAAGDKWSAFLEKQSTLAQMPLOETQNTLVKLOLQALQONSSVLSSEKSKRLNTIL 120  
 QY 121 NTMSTIYSTGKVCNPDNQECLLEPGINEIMANSLDYNERLWAMESRSRSEVGQOLRPLY 180  
 DB 121 NTMSTIYSTGKVCNPDNQECLLEPGINEIMANSLDYNERLWAMESRSRSEVGQOLRPLY 180  
 QY 181 EYVVLKEMARAHYEDYGDYWRGDEYVNGVDYDSRGQLLEDVHTTEETKPLYEHL 240  
 DB 181 EYVVLKEMARAHYEDYGDYWRGDEYVNGVDYDSRGQLLEDVHTTEETKPLYEHL 240  
 QY 241 HAYVRAKLMAVPSYISPIGCLPAHLIGDMWGRFTWNLVSLTVFPGQKPNIDVTDAVDAQ 300  
 DB 241 HAYVRAKLMAVPSYISPIGCLPAHLIGDMWGRFTWNLVSLTVFPGQKPNIDVTDAVDAQ 300  
 QY 301 AMDAQRIFKEAEKFFVSVGLPNNMTQGFWENSMLTDPGNVOKAVCHPTAMDLGKGRFRIIM 360  
 DB 301 AMDAQRIFKEAEKFFVSVGLPNNMTQGFWENSMLTDPGNVOKAVCHPTAMDLGKGRFRIIM 360

QY 361 CTKVTMDPEFLTAHHEMHTQYDMAYAAOPFLLRNCGNCFHEANGELMSLSAATPKHLKS 420  
 DB 361 CTKVTMDPEFLTAHHEMHTQYDMAYAAOPFLLRNCGNCFHEANGELMSLSAATPKHLKS 420  
 QY 421 IGLSPDFQEDNETELNFLKQALTYVGLTPTVYLKRWMMVTFGEIPKQMMKKWEM 480  
 DB 421 IGLSPDFQEDNETELNFLKQALTYVGLTPTVYLKRWMMVTFGEIPKQMMKKWEM 480  
 QY 481 KREIVGVPEVPHDETCDBPASFHVNSDYSFIRYTRTLXQFOFOALCOAAKHGEPH 540  
 DB 481 KREIVGVPEVPHDETCDBPASFHVNSDYSFIRYTRTLXQFOFOALCOAAKHGEPH 540  
 QY 541 KCDISNSTEAGOKL 554  
 DB 541 KCDISNSTEAGOKL 554

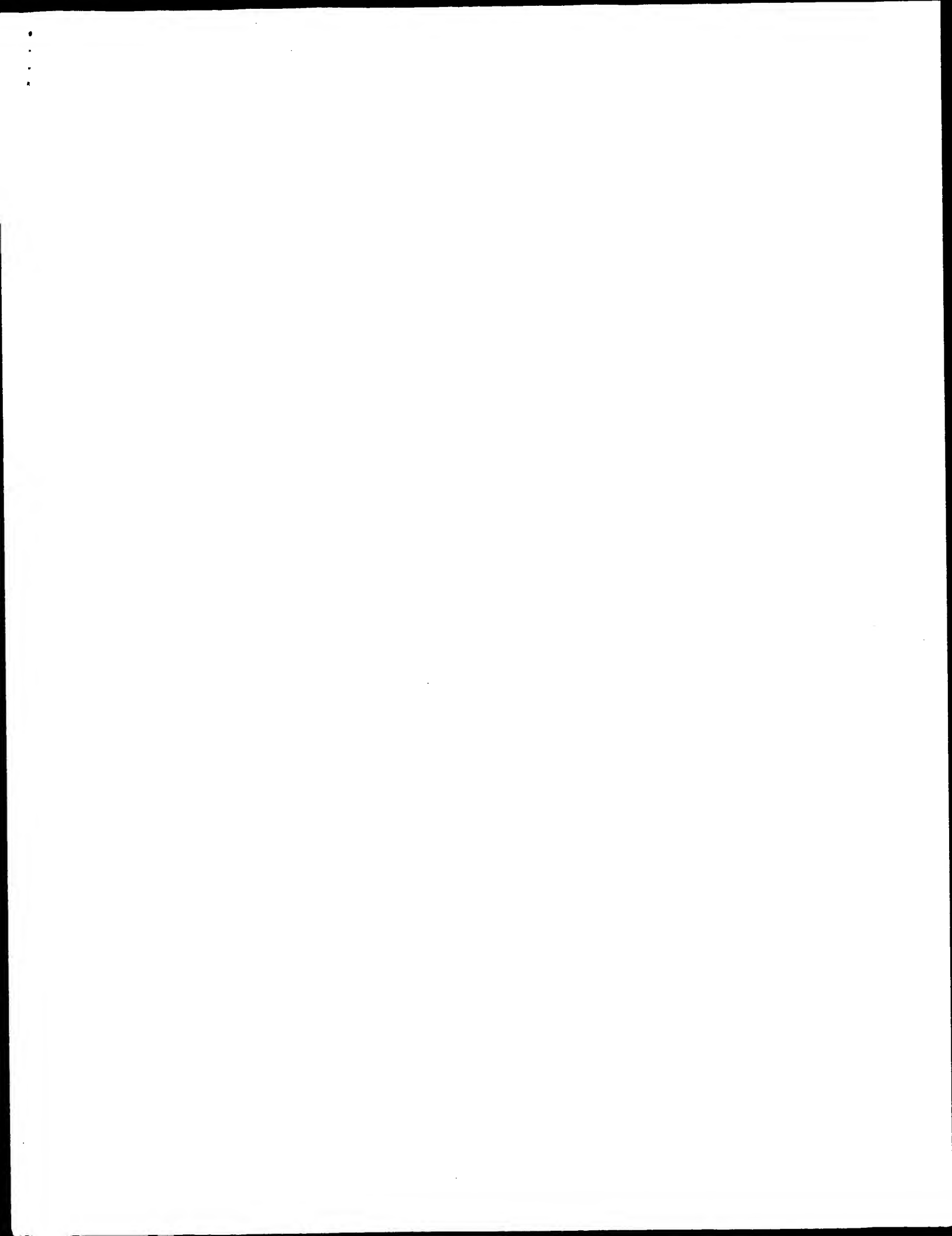
## RESULT 20

US-09-969-384-23  
 ; Sequence 23, Application US/09969384  
 ; Publication No. US20020192749A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Moore, et al.  
 ; TITLE OF INVENTION: Human Gene Polynucleotides, polypeptides, and Antibodies  
 ; FILE REFERENCE: pm055p1  
 ; CURRENT APPLICATION NUMBER: US/09/969,384  
 ; PRIOR FILING DATE: 2001-10-03  
 ; PRIOR APPLICATION NUMBER: PCT/US01/10542  
 ; PRIOR FILING DATE: 2001-04-02  
 ; PRIOR APPLICATION NUMBER: 60/236,384  
 ; PRIOR FILING DATE: 2000-09-29  
 ; PRIOR APPLICATION NUMBER: 60/194,118  
 ; PRIOR FILING DATE: 2000-04-03  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: Patentln Ver. 2.0  
 ; SEQ ID NO 23  
 ; LENGTH: 261  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-969-384-23

Query Match 31.7%; Score 1359; DB %: Length 261;  
 Best Local Similarity 99.6%; Pred. No. 3.1e-105;  
 Matches 252; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 62 MNNAAGDKWSAFLEKQSTLAQMPLOETQNTLVKLOLQALQONSSVLSSEKSKRLNTIL 121  
 DB 1 MNNAAGDKWSAFLEKQSTLAQMPLOETQNTLVKLOLQALQONSSVLSSEKSKRLNTIL 121  
 QY 122 TMSTIYSTGVCNPDNQECLLEPGINEIMANSLDYNERLWAMESRSRSEVGQOLRPLY 181  
 DB 122 TMSTIYSTGVCNPDNQECLLEPGINEIMANSLDYNERLWAMESRSRSEVGQOLRPLY 181  
 QY 182 EYVVLKEMARAHYEDYGDYWRGDEYVNGVDYDSRGQLLEDVHTTEETKPLYEHL 241  
 DB 182 EYVVLKEMARAHYEDYGDYWRGDEYVNGVDYDSRGQLLEDVHTTEETKPLYEHL 241  
 QY 242 AYVRAKLMAVPSYISPIGCLPAHLIGDMWGRFTWNLVSLTVFPGQKPNIDVTDAVDAQ 301  
 DB 242 AYVRAKLMAVPSYISPIGCLPAHLIGDMWGRFTWNLVSLTVFPGQKPNIDVTDAVDAQ 301  
 QY 302 WDAQRIFKEAEKFF 314  
 DB 302 WDAQRIFKEAEKFF 253

Search completed: March 13, 2003, 17:02:41  
 Job time : 25 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 7, 2003, 06:35:39 ; Search time 4862 Seconds  
(without alignments)  
4818.544 Million cell updates/sec

Title: US-09-978-385-2

Perfect score: 4291  
Sequence: 1 MSSSSWLLSLVAVTAQST.....ISKENNPQNTDVTQTSF 805

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 205640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlp  
-Q/cgn2.1/USPTO.spool/US09978385/runat.28022003.104716.1342/app-query.fasta.1.967  
-DB-GenEmbl -QEMT-fastap -SUPFIX-p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LIST=45  
-UNITS-bits -START=1 -END=1 -MATRIX-BLOSUM62 -TRANS-human40.cdi -LIST=45  
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20 -MODE=LOCAL  
-OUTFMT=pcio -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09978385.@cgn2.1.3745@runat.28022003.104716.1342 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_to:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hgt\_hum:\*  
31: em\_hgt\_inv:\*  
32: em\_hgt\_other:\*  
33: em\_hgt\_mus:\*  
34: em\_hgt\_pin:\*  
35: em\_hgt\_rod:\*  
36: em\_hgt\_mam:\*  
37: em\_hgt\_vrt:\*  
38: em\_sy:\*  
39: em\_hgt\_hum:\*  
40: em\_hgt\_mus:\*  
41: em\_hgt\_other:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4291	100.0	2415	6 AR135178	AR135178 Sequence
2	4291	100.0	2415	6 AX418984	AX418984 Sequence
3	4291	100.0	2415	6 E43987	E43987 ACE-analog
4	4291	100.0	2418	6 E39033	E39033 MPROT15 pol
5	4291	100.0	2599	6 E43988	E43988 ACE-analog
6	4291	100.0	3325	6 AF291820	AF291820 Homo sapi
7	4291	100.0	3334	6 AX047758	AX047758 Sequence
8	4291	100.0	3396	6 AR135177	AR135177 Sequence
9	4291	100.0	3396	6 AX418982	AX418982 Sequence
10	4291	100.0	3396	6 AX431513	AX431513 Sequence
11	4291	100.0	3405	6 AF241254	AF241254 Homo sapi
12	4288	99.9	3341	9 AB046569	AB046569 Homo sapi
13	4281	99.8	3341	6 AR135177	AR135177 Sequence
14	4142	96.5	3732	6 AX463938	AX463938 Sequence
15	3740.5	87.2	2262	6 E39034	E39034 MPROT15 pol
16	3579	83.4	2638	6 AX047762	AX047762 Sequence
17	3579	83.4	2739	10 BC026801	BC026801 Mus muscu
18	3562	83.0	2760	6 AB053181	AB053181 Mus muscu
19	3561	83.0	2638	6 AX047765	AX047765 Sequence
20	3509	81.8	2415	6 AX047764	AX047764 Sequence
21	2904	67.7	2415	6 BC034367	BC034367 Mus muscu
22	2167	50.5	4879	9 BC034367	BC034367 Mus muscu
23	1347.5	31.4	3681	9 BC034367	BC034367 Mus muscu
24	1347	31.4	2473	9 HSMACB	HSMACB
25	1344	31.3	2477	6 HSMACB	HSMACB
26	1344	31.3	2477	6 HSMACB	HSMACB
27	1344	31.3	2478	6 AR037213	AR037213 Sequence
28	1344	31.3	2478	6 AR137383	AR137383 Sequence
29	1337	31.2	4020	6 AR137383	AR137383 Sequence
30	1337	31.2	4020	6 HUMACB	HUMACB
31	1337	31.2	4022	6 AX429555	AX429555 Sequence
32	1337	31.2	4024	6 A00914	A00914 H. sapiens
33	1337	31.1	2418	10 MUSACB	MUSACB
34	1334	31.1	3813	10 MUSACB	MUSACB
35	1334	31.1	3939	6 AX147531	AX147531 Sequence
36	1334	31.1	4563	6 AX147503	AX147503 Sequence
37	1334	31.1	4694	10 MUSACB	MUSACB
38	1322	30.8	2551	9 HUMACB	HUMACB
39	1312	30.6	4050	5 CHKACB	CHKACB
40	1310	30.5	4014	10 AF201331	AF201331 Rattus norv
41	1310	30.5	4014	10 AF201332	AF201332 Rattus norv
42	1310	30.5	4014	10 RND03708	RND03708 Rattus norv
43	1310	30.5	4142	10 RND03734	RND03734 Rattus norv
44	1303	30.4	2490	9 HUMACB	HUMACB
45	1283.5	29.9	4803	4 OCANOE	OCANOE

RESULT 1

## ALIGNMENTS



[illegible]

```

|||||
Db 961 CCTAATATGACTCAAGATTCGGGAAATTCATGCTTACCGACCCAGAAATGTTCCAG 1020
Qy 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1021 AAGGACGCTGCTCCACAGCTTGGGACCTGGGAAAGGCGACTTCAGATTCCTTANG 1080
Qy 361 CysThrLysValIleMetAspPheLeuThrAlaHisIleGluMetGlyHisIleGln 380
Db 1081 TGCACAAGGTGACAAATGAGCAGCTCTGACAGCTCATCATGATGATGGGCAATCCAG 1140
Qy 381 TyrAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAspGlyLysGlnGlyPhe 400
Db 1141 TATGATATGGCATATCTGCAACCTTTCTCTAAGAAATGAGCTAATGAGGATTC 1200
Qy 401 HisGluAlaValGlyIleuIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1201 CATGACGCTTGGGAAATCATGCTCACTTTCGACGCCACACCTTAAGCATTTAAATCC 1260
Qy 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
Db 1261 ATTGCTCTCTGTCACCCGATTTTCAAGAAACAAATGAAACAAATTAACCTTCGCTC 1320
Qy 441 LysGlnAlaLeuThrIleValIleGlyThrLeuProPheThrTyrMetLeuGluIleTrpArg 460
Db 1321 AACCAAGCACTCACGATTTGGGACTCTGCCATTTACTTACATGTTAGGAAGTGGAGG 1380
Qy 461 TrpMetValPheLysGlyIleuIleProLysAspGlnTrpMetLysIleTrpGluMet 480
Db 1381 TGGATGCTCTTTAAAGGGGAAATTCCTCAAGACAGAGTGATGAAAGTGGTGGGAGATG 1440
Qy 481 LysArgGluIleValIleGlyValIleGluProValProHisAspGluThrTyrCysAspPro 500
Db 1441 AAGGAGAGATAGTGGGGGTGGGAGACCTGCTCCATATGAAATATCTGAGACCC 1500
Qy 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrpArgThrLeu 520
Db 1501 GCATCTCTGTTCCATGTTTCTTAAGATTAATCTATCTATGATATTAACAAGACCTT 1560
Qy 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis 540
Db 1561 TACCAATTCAGCTTCAAGAGCACTTGTCAAGCACCTTAACATGTAAGGCCCTCGCAC 1620
Qy 541 LysCysAspIleSerAsnSerThrGluAlaGlyIleLysLeuPheAsnMetLeuArgLeu 560
Db 1621 AAAAGTGCATCTCAACTCTACAGAGCTGGACAGAACTGTTCAATATAGCTGAGCCTT 1680
Qy 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValIleGlyAlaLysAsnMetAsn 580
Db 1681 GGAAATTCAGAACCTCGACCTGACCTTGGAATATGTTGTAGAGAGCAAAAGACATGAT 1740
Qy 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
Db 1741 GTAAGGCCACTGCTCAACTACTTGTAGCCCTTATTACTGCTGGCAAGACCAAGAACG 1800
Qy 601 AsnSerPheValIleTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
Db 1801 AATCTTTTGGGTGGAGTGAACGACTGGAGTGCATATGCAAGCCAAAGCATTAAGTG 1860
Qy 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640
Db 1861 AGGAAATCCCTAATAATGAGCTCTGAGATTAAGCATATGAAATGAAAGCAATGAATG 1920
Qy 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgIleTyrPheLeuLysValLysAsn 660
Db 1921 TACCTGTTCCGATCATCTGTGTGATATGCTATGAGGCGTACTTTTAAAGTAAATAAT 1980
Qy 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 1981 CAGATGATCTTTTGGGAGAGAGATGCGAGTGGCTAATTTGAAACCAAGATCTCC 2040
Qy 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
|||||

```

```

Db 2041 TTTAATTTCTTGTGCACCTGCACCTAAATAATGTGTCTGATATCATCTCTAGACCTGAGAT 2100
Qy 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2101 GAAAAGGCCATCAGAGATGTCCTCCGAGGCCGTATCATATATGCTTTCCTGTGAATGACAA 2160
Qy 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2161 AGCTTAGAGTTTCTTGGGATACACAGCAACCTTGACCTCTTACACAGCCCGCTTTCC 2220
Qy 741 IleTrpLeuIleValPheGlyValIleValMetGlyValIleValIleValIleLeu 760
Db 2221 ATATGCTGATGATGTTTGGAGTTGTGATGGAGATGATAGTGTGGCATTTGCATCTCTG 2280
Qy 761 IlePheThrGlyIleArgAspArgLysLysAsnLysAlaArgSerGlyGluAsnPro 780
Db 2281 AACTTCACTGCGATCAGATCAGATCGGAGAGAAATTAAGCAAGAGTGGAGAAATCTT 2340
Qy 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800
Db 2341 TATGCTCATCGATATTTAGCAAGAGAAATATATCCAGATTCCTCAAAACACTGATGAT 2400
Qy 801 ValGlnThrSerPhe 805
Db 2401 GTTCAGACCTCCTTT 2415

```

## RESULT 3

```

E43987 LOCUS E43987 2415 bp DNA linear PAT 31-JAN-2002
DEFINITION ACE-analogous gene.
ACCESSION E43987
VERSION E43987.1 GI:18629190
KEYWORDS JP 2001046072-A/1.
SOURCE unidentified.
ORGANISM unidentified.

```

```

REFERENCE 1 (bases 1 to 2415)
AUTHORS Sugano,S. and Komatsu,T.
TITLE ACE-analogous gene
JOURNAL Patient: JP 2001046072-A 1 20-FEB-2001;
OTSUKA PHARMACEUT CO LTD

```

## COMMENT

```

OS Unknown
PN JP 2001046072-A/1
PD 20-FEB-2001
PE 06-AUG-1999 JP 1999223892
PR
PI SUMIO SUGANO, TAKAMI KOMATSU
PC C12N15/09, A61K31/00, A61K31/7088, A61K38/00, A61K38/55,
A61K39/395, PC A61K39/395,
PC A61K39/395, A61K48/00, A61P9/12, C07K14/47, C07K16/08, C12N1/15, PC
C12N1/19,
PC C12N1/21, C12N5/10, C12Q1/68, G01N33/53, C12N15/00, A61K37/02, PC
A61K37/64,
PC C12N5/00
CC
FH Key 1. .2415 Location/Qualifiers
FT source 1. .2415
FT /organism='Unknown'.

```

## FEATURES

```

source 1. .2415 Location/Qualifiers
BASE COUNT 743 a 484 c 554 g 634 t
ORIGIN

```

## Alignment Scores:

```

Pred. No.: 0 Length: 2415
Score: 4291.00 Matches: 805
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```





QY 721 SerLeuGIuPheLeuGIyIIeGIInProThLeuGIyProProAsnGIuProProValSer 740  
 Db 2161 AGCCTAGAGTTCTGGGGGATACAGCCAACTGGACCTCTTAACAGCCCTGTTCC 2220  
 QY 741 IIeTrpLeuIIeValPheGIyValIleMetGIyValIIeValIleGIyIIeValIleLeu 760  
 Db 2221 AATAGGCTGATGTTTGGAGTTGGATGGAGTGAATGGTGGTGGCATGTCACTCG 2280  
 QY 761 IIePheThrGIyIIeLeuArgAspArgIySIySIyAsnIySIyAlaArgSerGIyIuAsnPro 780  
 Db 2281 ATCTTCACGTGGATCAGAGATCGGAGAGAAAATAAAGCAAGAGAGTGGAAATCTCT 2340  
 QY 781 TyrAlaSerIIeAspIIeSerIySIyGIyIuAsnAsnProGIyPheGIuAsnThraSpAsp 800  
 Db 2341 TATGCCCTCATCATATATTAGCAAGAGAAATAATCATCAGGATTCCAAACACTGATGAT 2400  
 QY 801 ValGIuThraSerPhe 805  
 Db 2401 GTTCAGACCTCTCTT 2415  
 RESULT 4  
 E39033  
 LOCUS E39033 2418 bp DNA linear PAT 18-JUN-2001  
 DEFINITION MPROT15 polypeptide and MPROT15 polynucleotide.  
 ACCESSION E39033  
 VERSION E39033.1 GI:13017695  
 KEYWORDS JP 1999318472-A/1.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 2418)  
 AUTHORS Christopher D.S. and Nicola B.  
 TITLE MPROT15 polypeptide and MPROT15 polynucleotide  
 JOURNAL Patent: JP 1999318472-A 1 24-NOV-1999.  
 COMMENT SMITHKLINE BEECHAM CORP PUBLIC LTD CO  
 OS Homo sapiens (human)  
 PN JP 1999318472-A/1  
 PD 24-NOV-1999  
 PE 22-JAN-1999 JP 1999014949  
 PR 13-MAY-1998 GB 9810373.2, 18-AUG-1998 GB 9818009.4 PI  
 CHRISTOPHER D SAZAN, NICOLA BAGESU  
 PC C12N15/09, A61K31/00, A61K31/00, A61K31/00, A61K31/00, A61K31/00,  
 PC A61K38/00,  
 PC A61K39/395, A61K45/00, A61K48/00, C07K14/47, C12P21/02, C12Q1/68,  
 PC G01N33/50, G01N33/50, G01N33/566, G01N33/68//C12P21/02,  
 C12R1:91, PC C12N15/00,  
 PC A61K37/02  
 CC  
 FH  
 FT  
 KEY Location/Qualifiers  
 source 1..2418  
 Location/Qualifiers  
 1..2418  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 744 a 484 c 555 g 635 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 2418  
 Score: 4291.00 Matches: 805  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 DB: 6  
 US-09-978-385-2 (1-805) x E39033 (1-2418)  
 QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValIhAlaIaGInSerThr 20  
 Db 1 AITGTCAAGCTCTTCTGGCTCTTCACGCTTGTGGTGTGAATCTGCTCAAGTCCACC 60

QY 21 ILeGIuGIuAlaIaIySIyThrPheLeuAspIySIyPheAsnIySIyAlaIaGIuAspLeuPhe 40  
 Db 61 ATTGAGAAACAGGCCCAAGACATTTTGGACAGAGTTTAAACCAAGCAAGCGAGACTGTTC 120  
 QY 41 TyrGIuInSerSerLeuAlaSerTrpAsnTyrAsnThraSnIIeThrGIuGIuAsnValGIu 60  
 Db 121 TATCAAGTTCACCTTCTCTTGGAAATATTAACCAATATTAATCAAGAAAGTGTCAA 180  
 QY 61 AsnMetAsnAspAlaGIyAspIySIyTrpSerIhPheLeuSIyGIuInSerThrLeuAla 80  
 Db 181 AACATGAATTAATGCTGGGGACAAATGCTCTCTTTTAAAGGAACAGTCCACCTTGCC 240  
 QY 81 GInMeTyrProLeuGIuGIuIIeGIuAsnLeuThraValIySIyLeuGIuInLeuAlaLeu 100  
 Db 241 CAATGTATCCACTCAAGAAATTCAGATATTCACAGTCAAGCTTCACCTGACGCTCTT 300  
 QY 101 GIuGIuAsnGIySerSerValIeUsnSerGIuAspIySIySerIySIyArgLeuAsnThrIleu 120  
 Db 301 CAGCAAAATGGGTCTTCAGTCTCTCAGAAAGCAAGCAAAACGTTGAACACAAATCTTA 360  
 QY 121 AsnThraMetSerThrIIeTyrSerThrGIySIyValCysAsnProAspAsnProGIuGIu 140  
 Db 361 AATACAAATGAGCACCATCTACAGTACAGTGAAGATTTGTATCCAGATTAATCCACAGAA 420  
 QY 141 CysLeuLeuLeuGIuProGIyLeuAsnGIuIleMetAlaAsnSerLeuAspTyrAsnGIu 160  
 Db 421 TCTATTACTTACGTGAACCAAGCTTGAATGAATTAATGGCAAAACATTTAGACTACAAATGAG 480  
 QY 161 ArgLeuTrpAlaTrpGIuSerTrpArgSerGIuValGIySIyGIuLeuArgProLeuTyr 180  
 Db 481 AGGCTCTGGGCTTGGGAAAGCTGAGATCGAGTGGGCAAGCAGTGAAGCCATTTAT 540  
 QY 181 GIuGIuTyrValValIeLeuIySIyAsnGIuMetAlaArgAlaAsnIySIyGIuAspTyrGIy 200  
 Db 541 GAAGAGTATGTGGCTTAAAAAATGAGATGGCAAGCAAGCAAAATCTTATGAGACTATGGG 600  
 QY 201 AspTyrTrpArgIyAspTyrGIyValAsnGIyValAspGIyTyrAspTyrSerArgGIy 220  
 Db 601 GATTATTGGAGAGAGACATATGAAGTAAATGGGGTAGTGGCTTATGACTACAGCGCGGC 660  
 QY 221 GInLeuIIeGIuAspValGIuInSIyThrPheGIuGIuIleIySIyProLeuTyrGIuInSIyLeu 240  
 Db 661 CAGTGTATGAAGATGTGGAACATATCTTTGAAGAAGATTAACCATATATGAACATCTT 720  
 QY 241 HisAlaTyrValAlaArgAlaIySIyLeuMetAsnAlaTyrProSerTyrIleSerProIleGIy 260  
 Db 721 CATGCTATGTGAGGGCAAAAGTGAATGATGATCTATCTCTCTATATCAGTCCAAATGGA 780  
 QY 261 CysLeuProAlaHisIySIyLeuGIyAspMetTrpGIyArgPheTrpIhAsnLeuTyrSer 280  
 Db 781 TGGCTCCCTGCTCATTTCTCTTGGATATATGGGGTAGATTTTGGCAAAATCTGTACTCT 840  
 QY 281 LeuThraValProPheGIyGIuIySIyProAsnIIeAspValThraSpAlaMetValAspGIu 300  
 Db 841 TTGACAGTTCCTTGGGACAAACCAACATATATCTTACTGATCAAAATGGTGGACAG 900  
 QY 301 AlaTrpAspAlaGIuInSIyIlePheIySIyGIuAlaIySIyPhePheValSerValGIyLeu 320  
 Db 901 GCCTGGATGCAACAGCATATTCAGAGAGGCGAGAAAGTCTTGTATCTGTTGGCTT 960  
 QY 321 ProAsnMetTrpGIuInSIyPheTrpGIuAsnSerMetLeuThraSpProGIyAsnValGIu 340  
 Db 961 CCTATATGACTCAAGAGTCTTGGGAAATTCATGCTAACGAGCCACAGAAATGTTCCAG 1020  
 QY 341 LysAlaValCysHisProThAlaTrpAspLeuGIySIyGIyAspPheArgIIeLeuMet 360  
 Db 1021 AAAGAGTCTCCATCCACAGCTTGGGACCTTGGGAGAGGCGCACTTCAGAGATCTTATG 1080  
 QY 361 CysThrIySIyValIhPheMetAspAspPheLeuThraHisnIySIyGInMetGIyHisIleGIu 380  
 Db 1081 TGCACAAAGGTGACAAATGAGAGACTCTCTGACAGCTCATATGATGAGTGGGCAATTCAG 1140

QY 381 TyrAspMetAlaTyrAlaIaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPhe 400  
 |||||||  
 Db 1141 TATCATATGGCATATGCTGCACAAACCTTTCTCTAGAAATGAGCATATGAAAGATTTC 1200  
 QY 401 HisGluAlaValGlyIleuMetSerLeuSerAlaAlaThrPolysHisLeuYsser 420  
 |||||||  
 Db 1201 CATTAAGCTGTGGGAAATCATCTACTTCTGCACACCACCTAAGCATTTTAAATTC 1260  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440  
 |||||||  
 Db 1261 ATTTGCTCTTCTGTCACCCGATTTTCAAGAGACATGATAAACAAGAAATCTTCTCTC 1320  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrIleMetLeuGluYsserArg 460  
 |||||||  
 Db 1321 AAACAGACCTCAGCATTTGGGAGCTGCGCATTTACTTACATGTTAGAGAACTGGAGG 1380  
 QY 461 TrpMetValPheYssGlyGluIleProLysAspGlnThrPheLysLysTrpTrpGluMet 480  
 |||||||  
 Db 1381 TGGATGGCTTTTAAAGGGGAATTCACAAAGACAGTGAAGAAAGTGGTGGAGATG 1440  
 QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTrpCysAspPro 500  
 |||||||  
 Db 1441 AAGCGAGATAGTGGGGTGGTGAACCTGTGCCCATATGAACATACTGTGACCC 1500  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrpThrArgThrLeu 520  
 |||||||  
 Db 1501 GCATCTCTGTCATGTTTCTTAATGATTTACTTACTTTCATTTACACAGAACCTT 1560  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlnAlaLeuGlnAlaAlaYssHisGlnGlyProLeuHis 540  
 |||||||  
 Db 1561 TACCAATTCAGATTTCAGAAAGACACTTGTTCAGACAGCTAAACATGAAGGCCCTCTGCAC 1620  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlnGlnYssPheAsnMetLeuArgLeu 560  
 |||||||  
 Db 1621 AAATGTGATCTCACTCAACTGACAGAGCTGACAGAACTGTCATATATCTCTGAGGCTT 1680  
 QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGlnAsnValAlaGlyAlaYssMetAsn 580  
 |||||||  
 Db 1681 GGAAATATCAGAACCTTGACCTTACATTTGGAAATGTTTGAAGAGCAAAAGACATGAT 1740  
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpPheLysAspGlnAsnLys 600  
 |||||||  
 Db 1741 GTAAAGCCACTGCTCACTACTTGTAGGCCCTTATTTACCTGCTGAAAGACAGAACAG 1800  
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620  
 |||||||  
 Db 1801 AATCTTTTGGTGGATGGATGACCGACCTGAGTCCATATGACGACCAAGCATCAAGT 1860  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlnLysPlyAlaTyrGluTrpAsnAspAsnGluMet 640  
 |||||||  
 Db 1861 AGGATTAAGCTTAATACAGCTCTTGTGAGATTAAGCATATGATGAAGACATGAATG 1920  
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660  
 |||||||  
 Db 1921 TACTGTTCCGATCATCTGTTGATATGATGATGAGGAGTACTTTTAAAGTAAATAAT 1980  
 QY 661 GlnMetIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 |||||||  
 Db 1981 CAGATGATCTTTTGGGAGGAGATGCGAGTGGCTAATTTTAACCAAGAAATCTTCC 2040  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700  
 |||||||  
 Db 2041 TTTAATTTCTTTGTCTACCTCAGCTAAATATGTTGATATCATCTTGAAGACTGAAGT 2100  
 QY 701 GlnLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 |||||||  
 Db 2101 GAAAGAGCATCAGATGTCGCGAGCCCTATCAATGATGATCTTCCGCTTGAATACAAAC 2160  
 QY 721 SerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 |||||||  
 Db 2161 AGCTTAGAGTTCTTGGGGATACAGCCACACACTTGGACCTCTTAACAGGCCCTGTGTTCC 2220  
 QY 741 IleTrpLeuIleValPheGlyValAlaMetGlyValIleValIleValIleValIleLeu 760

Db 2221 ATATGGCTATGTTTGTGGATTTGATGGAGTGAATAGTGTGCTGATGTCATCTG 2280  
 |||||||  
 QY 761 IlePheThrGlyIleArgAspArgLysLysAsnLysAlaArgSerGlyGluAsnPro 780  
 |||||||  
 Db 2281 ATCTTCACCTGATCGATATGCGAAGAGAAATTAAGCAAGAGTGGAGAAATATCT 2340  
 QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800  
 |||||||  
 Db 2341 TATGCTTCATGATATTTGCAAGAGAGAAATATATCCGATTTCCAAACAGCTGATGAT 2400  
 |||||||  
 QY 801 ValGlnThrSerPhe 805  
 |||||||  
 Db 2401 GTTCAGACCTCCTTT 2415  
 |||||||  
 RESULT 5  
 E43988  
 LOCUS E43988 2599 bp DNA linear PAT 31-JAN-2002  
 DEFINITION ACE-analogous gene.  
 ACCESSION E43988  
 VERSION E43988.1 GI:18629191  
 KEYWORDS JP 2001046072-A/2.  
 SOURCE unidentified.  
 ORGANISM unidentified.  
 REFERENCE 1 (bases 1 to 2599)  
 AUTHORS Sugano, S. and Komatsu, T.  
 TITLE ACE-analogous gene  
 JOURNAL Patent: JP 2001046072-A 2 20-FEB-2001;  
 OTSUKA PHARMACEUT CO LTD  
 OS Unknown  
 PN JP 2001046072-A/2  
 PD 20-FEB-2001  
 PE 06-AUG-1999 JP 1999223892  
 PR  
 PL SUMIO, SUGANO, TAKAMI, KOMATSU  
 PC C12N15/09, A61K31/7088, A61K38/00, A61K38/55,  
 A61K39/395, PC A61K39/395,  
 PC A61K39/395, A61K48/00, A61P9/12, C07K14/47, C07K16/08, C12N1/15, PC  
 C12N1/19,  
 PC C12N1/21, C12N5/10, C12Q1/68, G01N33/53, C12N15/00, A61K37/02, PC  
 A61K37/64,  
 PC C12N5/00  
 CC  
 FH Key  
 FT CDS Location/Qualifiers  
 FEATURES  
 source 1..2599 Location/Qualifiers  
 /organism="unidentified"  
 /db\_xref="taxon:32644"  
 BASE COUNT 807 a 503 c 591 g 698 t  
 ORIGIN  
 Alignment Scores:  
 Pred. NO.: 0 Length: 2599  
 Score: 4291.00 Matches: 805  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-978-385-2 (1-805) x E43988 (1-2599)  
 QY 1 MetSerSerSerTrpLeuLeuLeuSerLeuValAlaValThrAlaIaGlnSerThr 20  
 |||||||  
 Db 55 ATGTCAACCTTCTTCTGCTCTCTCAGCCTTGTCTGTAACGTGCTCAGTCCACC 114  
 |||||||  
 QY 21 IleGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40  
 |||||||  
 Db 115 ATTGAGAAAGGCCAAGACATTTTGGACAAAGTTTAACACGAGAAAGCCGAGACCTGTTCC 174  
 |||||||  
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrTrpAsnThrAsnIleThrGlnGluAsnValGln 60  
 |||||||

Db 175 TATCAAGTCTACTGCTTCTTGGAATTAACACCAATATTAAGTCA 234  
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuYsgLugInserThrLeuAla 80  
 Db 235 AACATGATATATGCTGGGCAATATGCTGCTTTTAAAGACAGCTCCACTTGGC 294  
 QY 81 GluMetLysTrpLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100  
 Db 295 CAAATGATACCTACACAGAAATTCAGAAATTCACAGTCAAGCTCAGCTGAGCTCTT 354  
 QY 101 GlnGlnAsnGlySerValLeuSerGlnAspLysTrpLysArgLeuAsnThrIleLeu 120  
 Db 355 CACGAAATATGGCTCTCAGTCTCTCAGAAAGACAGACAAAGGTTGAACCAATTCCTA 414  
 QY 121 AsnThrMetSerThrIleLysSerThrGlnLysValLysAsnProAspAsnProGlnGlu 140  
 Db 415 AATACATATGAGCACCATCTACAGTACAGTGAAGAAAGTTGTAACCCAGATATCCACAGAA 474  
 QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTrpAsnGlu 160  
 Db 475 TGTCTTATCTTGAACCAAGGTTGAATGAATATGCAACAGTTTACACTACATGAG 534  
 QY 161 ArgLeuTrpAlaTrpLysSerTrpArgSerGlnValGlyLysGlnLeuAsnProLeuTrp 180  
 Db 535 AGGCTCTGGGCTGGGCAAGCTGGAGATCTGAGCTGGCAGACAGCTGAGGCTATTAAT 594  
 QY 181 GluGluTrpValValLeuLysAsnGluMetAlaArgAlaAsnIleTyrosLysProLysGly 200  
 Db 595 GAAGATATATGCTGCTTGAAGAAATGAGTGGCAAGAGCAATCATATATGAGCATATGG 654  
 QY 201 AspTrpTrpArgGlyAspTrpGlnValAsnGlyValAspGlyTrpAspTrpSerArgGly 220  
 Db 655 GATATATGAGAGAGACATATGATAAATGGGGTGAAGTGGCTATGACTACAGCCGCGC 714  
 QY 221 GlnLeuIleGlnAspValGluIleThrPheGlnGluIleLysProLeuTrpGlnIleLeu 240  
 Db 715 CAGTGAATGAGAGATGAGACATACCTTGAAGAGATTAACCTTTATATGACATCTT 774  
 QY 241 HisAlaTrpValAlaArgAlaLysLeuMetAsnAlaTrpProSerTrpIleSerProIleGly 260  
 Db 775 CATGCTATATGAGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 834  
 QY 261 CysLeuProAlaHisLeuLeuGlnLysPheTrpGlyArgPheTrpThrAsnLeuTrpSer 280  
 Db 835 TGCCTCCCTGCTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894  
 QY 281 LeuTrpValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 Db 895 TTGACAGTTCCTTTGGACAGAAACCAACATAGATGTACTGATGCAATGTGGACAG 954  
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerValGlyLeu 320  
 Db 955 GCGTGGATGACAGACAGAAATATTCAGAGAGCCGAGAAAGTCTTTGATCTGGTGGCTT 1014  
 QY 321 ProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsnValGln 340  
 Db 1015 CCTATATAGACTCAAGAGATCTCGGAAATTCATGCTAACGAGCCAGGAAATGTTCCAG 1074  
 QY 341 LysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 Db 1075 AAAGCAGTCTGCAATCCACAGCTTGGAGACCTGGGGAAGGCGCACTTCAGATCCCTATG 1134  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlnGluMetGlyHisIleGln 380  
 Db 1135 TGCAACAAGGTGACATGAGAGCTCTCTGACAGCTCATATGATGAGGGGCAATTCAG 1194  
 QY 381 TyrAspMetAlaTrpAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPhe 400  
 Db 1195 TATATATATGCAATCTGCAACAACCTTTCTGCTAAGAAATGAGAGCTATATGAGATTC 1254  
 QY 401 HisGlnAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
 Db 1255 CATACAGCTGTGGGAAATCATGCTCTTTCGACCCACACCTTAAGCATTTAAATCC 1314  
 QY 421 IleGlyLeuLeuSerProAspPheGlnLysAsnGluTrpGluIleAsnPheLeu 440  
 Db 1315 AATGCTCTTCTGACCCGATTTTTCAGAGACATATGAAGAAATTAACCTCTGCTC 1374  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheTrpThrMetLeuGlnLysTrpArg 460  
 Db 1375 AATCAAGACATCAGCATTTGTTGGACCTGCGCATTTACTTACATGTTAGAAATGGAGG 1434  
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpGlnMet 480  
 Db 1435 TGGATGCTCTTAAAGGGGAATTCGCCAAAGACAGTGGATGAAAGTGGTGGAGATG 1494  
 QY 481 LysArgGluIleValGlyValValGluProValProHisArgGluThrTrpCysAspPro 500  
 Db 1495 AAGCAGAGATATGTTGGGCTGCTGAGACCTTGGCCCATGATGAAACATCTGTGACCCC 1554  
 QY 501 AlaSerLeuPheHisValSerAsnAspTrpSerPheIleArgTrpTrpThrTrpLeu 520  
 Db 1555 GCATCTCTGTTCCATGTTCTTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1614  
 QY 521 TyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlnGluProLeuHis 540  
 Db 1615 TACCAATTCAGTTTCAAGAAAGCACTTGTCAACAGCACTAAACATGAAGGCTCTGAC 1674  
 QY 541 LysCysAspIleSerAsnSerThrGlnAlaGlyLysLeuLeuPheAsnMetLeuGluLeu 560  
 Db 1675 AATATGACATCTTAACATCTACAGAGAGTGGACAGAACTGTTCATATATGCTGAGGCTT 1734  
 QY 561 GlyLysSerGlnProTrpThrLeuAlaLeuGlnAlaAsnValValGlyLysAsnMetAsn 580  
 Db 1735 GGAATATCAGAACCCCTGAGCCCTGATGGAATATGTTCTGAGACCAAGAAATCATGAT 1794  
 QY 581 ValArgProLeuLeuAsnTrpPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLys 600  
 Db 1795 GTAAGGCCATGCTCAACTCTTGAAGCCCTTATTAACGCTGGAAGACCAAGCAAG 1854  
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProValAlaAspIleSerIleLysVal 620  
 Db 1855 AATCTTTTGGGATGAGATTAACGATGAGAGCTGAGAGCTGATATGACCAAGACATCAAGT 1914  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrosLysProAsnAspAsnGluMet 640  
 Db 1915 AGGATTAAGCCCTAAATCAGCTCTTGGAGATTAAGCATATGAAATGAAACATGAAATG 1974  
 QY 641 TyrLeuPheArgSerSerValAlaTyrosAlaMetArgGlnThrPheLeuLysValLysAsn 660  
 Db 1975 TACCTGTCGATCATCTGTTCATATGCTATGAGGAGAGTACTTTTAAAGTAAATAAT 2034  
 QY 661 GluMetIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 Db 2035 CAGATATCTTTTGGGAGAGAGATGTCGAGTGGCTTAATTTGAACCAAGATCTCC 2094  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTrpGluVal 700  
 Db 2095 TTAAATTTCTTCTGCTACCTCAGCTTAATAATGTTGATGATATCTCTAAGACTAAGTT 2154  
 QY 701 GluLysAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 Db 2155 GAAAGGCATCAGATATGCTCCGAGAGCTATATCATATGATGCTTCCGCTGATATGACAC 2214  
 QY 721 SerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 Db 2215 AGCTGTAGATTTCTGGGATACAGCCACACTTGGACCTCTTAACACACCCCTCTTTCC 2274  
 QY 741 IleThrLeuIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeu 760  
 Db 2275 AATATGCTGATGTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2334  
 QY 761 IlePheThrGlyIleArgAspArgLysLysAsnLysAlaArgSerGlyGlnAspPro 780  
 Db 2335 ATCTTACAGGATCAGAGATGAGAGAAAGAAATTAACCAAGAGTGGAGAAATCTT 2394

QY 781 Tyra1aSer1leasp1leSer1ySg1yLuasnaSnaProG1yPheG1nAsnThrAspAsp 800  
 Db 2395 TATCCTTCACATGCATATATAGCAAGAAATATATCCAGATTCCAAAACACTGATGAT 2454  
 QY 801 ValGlnThrSerPhe 805  
 Db 2455 GTTCAGACCTCTCTT 2469

RESULT 6  
 AF291820 3325 bp mRNA linear PRI 15-JUN-2001  
 LOCUS Homo sapiens ACE-related carboxypeptidase ACE2 mRNA, complete cds.  
 DEFINITION AF291820  
 ACCESSION AF291820.1 GI:9802432  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 3325)  
 Donoghue, M., Hsieh, F., Baronas, E., Godbout, K., Gosselin, M.,  
 Stagliano, N., Donovan, M., Woolf, B., Robison, K., Jeyaseelan, R.,  
 Breitbart, R.E. and Acton, S.  
 A novel angiotensin-converting enzyme-related carboxypeptidase  
 (ACE2) converts angiotensin I to angiotensin 1-9  
 Circ. Res. 87 (5), E1-E9 (2000)

JOURNAL MEDLINE 20429895  
 PUBMED 10969042  
 REFERENCE 2 (bases 1 to 3325)  
 Donoghue, M., Woolf, B., Robison, K. and Acton, S.  
 Direct Submission  
 Submitted (01-AUG-2000) Cardiovascular Biology, Millennium  
 Pharmaceuticals, Inc., 75 Sidney Street, Cambridge, MA 02139, USA

FEATURES  
 source location/Qualifiers  
 1..3325  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 21..2438  
 /codon\_start=1  
 /product="ACE-related carboxypeptidase ACE2"  
 /protein\_id="AAF99721.1"  
 /db\_xref="GI:9802432"  
 /translation="MSSSSWLLSLVAVTAASSTIEEQAKTFIDKFNHEADLFYQSS  
 LASWYNTNITEENYONMNNAGDKWSAFIEKQSTIAOMYPLQEIONTLVKIOALQO  
 NGSSVLEDKSKRLNTILNTSTYSGTQCPDMPQECILIEPLNTEMANSLDYNE  
 RLMANESMSPEYKQRLREYEVYVKKEMARANTEDIGDWRGDYEVNGVDYDS  
 RQGLIEDVEHTETKPLYEHLHAYVRAKLMAYPSYISPIECPLAHLIGEMWGRMT  
 NLYSLVPEGQKPNIDVIDAMVADQAKLFKEKEFEVSGLPMTQGWENSMLE  
 DPGVQKAVCHPTAMDLDGKDFRILMCTKRYMDDELTAHHEMGILOYMAYAAOPEL  
 RANGANGFEHAEVGEIMSLSAATPKHLKSLGLSPFOEDNETETFLKOLATTVGTL  
 PPTYLEKRWKRVFKETIPKQDMKKMKMKREYGVYEPVPHDTYCDPASPISVSN  
 DSYTRYTRTYIYQFOEALCOAKKHGEPFLKCDISNSTEAGQKLFNMLRIGSEPM  
 TTALENVGAKKMKNVRPLNTEFPLETWLKDQNKSEFGWSMDSPYADQSLKVALISL  
 KSLADKAIEMWDMENMYLERSSVAYAMQYELKYNQMILEGEDVRYANLKPISFN  
 FVTAPKNSDIIIPTEVEKAIKMSRSRINQALFRLNDMSLELGIQPLIGPNOPVS  
 IMLIEGVVGVGIYVGIYILFTGIRDKKKKKKNSGENPYASIDISGENNPGQNT  
 DDVQTSF"

BASE COUNT 1022 a 637 c 747 g 919 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0 Length: 3325  
 Score: 4291.00 Matches: 805  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-978-385-2 (1-805) x AF291820 (1-3325)

QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 21 ANGTCAAGCTCTTCTCGGCTCTCTCAGCCTTGTGCTGTAAGTCTGCTACGCCACC 80  
 QY 21 TleG1uG1nAla1a1ySThrPheLeuAsp1ySpheAsnH1eG1uAlaG1uAspLeuPhe 40  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 81 ATTGAGAACAGAGCCAGACATTTTGGACAGATTAAACAGAGAGCCGAGACCTGTTC 140  
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTrpAsnThrAsnThrGlnGluAsnValGln 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 141 TATCAAGTTCACCTGCTTCTTGGAATTAATACCAATATTATACGAAGAAATGCCAA 200  
 QY 61 AsnMetAsnAlaG1yAsp1yStrSerAlaPheLeu1ySg1uGlnSerThrLeuAla 80  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 201 AACATGAATAAATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCCACTTCC 260  
 QY 81 GlnMetTrpProLeuGlnGlu1leGlnAsnLeuThrVal1ySLeuGlnAlaLeu 100  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 261 CAATATGATCCACTACACAAATTCAGATTCACAGTCAAGCTTCAGCTGAGGCTCTT 320  
 QY 101 GlnGlnAsnG1ySerSerValLeuSerG1uAsp1ySeryArgLeuAsnThr1leLeu 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 321 CAGCAAAATGGGTCTTCAAGTCTCTCAGAGACAAAGACGGTTAACAACAATTTCTA 380  
 QY 121 AsnThrMetSerThr1leTyrSerThrG1yVal1yCysAsnProAspAsnProGlnGlu 140  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 381 AATACAAATAGACACCATCATGACTGGAAGAAAGTTTGAACCCGATTAATCCACAA 440  
 QY 141 CysLeuLeuGlnProG1yLeuAsnGlu1leMet1aAsnSerLeuAspTrpAsnGlu 160  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 441 TGCCTATTACTTGAACCAAGGTTGATGAATAAATGAGCAACAGTTTGAATACATAG 500  
 QY 161 ArgLeuThrAlaTrpGlnSerTrpArgSerG1uVal1yG1ySglnLeuArgProLeuTyr 180  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 501 AAGCTCTGGGCTTGGAAGGCTGAGATCTGAGGTCGACAGACCTGAGCCATTAAT 560  
 QY 181 GlnGluTyrVal1leuLeu1ySAsnG1uMet1aArgAlaAsnH1eTyrG1yAspTrpG1y 200  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 561 GAAGAGTATGCTGCTTGAAGAAATGAGATGGAAGCAAGCAATCATATAGAGACTATGG 620  
 QY 201 AspTrpTrpArgG1yAspTrpG1uVal1yAsnG1yVal1yAspG1yTyrAspTrpSerArgG1y 220  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 621 GATTAATGGAGAGGAGCTATGAAATGAGGAGTATGCTATGATACAGCCGCGGC 680  
 QY 221 GlnLeu1leG1uAspValGlnH1eThrPheG1uGln1leTyrProLeuTyrGlnH1eLeu 240  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 681 CAGTTGATTAAGAGATGAGACATACCTTTGAAGAGATTAAACCAATTAATGAACATCTT 740  
 QY 241 H1eAlaTyrVal1ArgAla1ySLeuMetAsnAlaTyrProSerTyr1leSerPro1leG1y 260  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 741 CATGCCATATGAGGCAAGTGAATGCTATCTCTCTATATACAGTCCAAATTCGA 800  
 QY 261 CysLeuProAlaH1eLeuLeuG1yAspMetTrpG1yArgPheTrpThrAsnLeuTyrSer 280  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 801 TGCCTCCCTGCTCAATTCCTTGATATGTGGGGTGAATTTTGGACAAATCTGTACTCT 860  
 QY 281 LeuThrValProPheG1yGln1ySProAsn1leAspVal1ThrAspAlaMetVal1ySgln 300  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 861 TTGACAGTTCCTTTTGGACAGAAACCAACATAGTGTACTGTAGTGCATGTGGACAG 920  
 QY 301 AlaTrpAspAlaGlnArg1lePhe1ySg1uAlaG1u1ySPhenVal1SerVal1yG1uLeu 320  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 921 GCCTGGATGCACAGAGAAATATTCAGAGAGGCCGAGAAATCTTTGATATCGTTGGCTTT 980  
 QY 321 ProAsnMetTrpGlnG1yPheTrpG1uAsnSerMetLeuThrAspProG1yAsnValGln 340  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 981 CCTAATATGACTCAAGATTCCTGGGAAATTCATGTCAAGGACCCAGCAAAATGTTTCAG 1040  
 QY 341 LysAlaVal1ySAsnProThrAlaTrpAspLeuG1y1ySg1yAspPheArg1leLeuMet 360  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1041 AAAGAGTCTGCTATCCACAGCTTGGGACCTGGGGAAGGGCGACTTCAGAGATCTTATG 1100  
 QY 361 CysThrLysVal1ThrMetAspAspPheLeuThrAlaH1eS1yGlnMetG1yH1eGln 380  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1101 TGCAAAAGGTGACAAATGAGACACTCTCTGACACCTCATCATGTGAGATGGGCAATATCCAG 1160

QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyPhe 400  
 |||||||  
 Db 1161 TATGATATGCGATATGCGCAACCTTTCTCTAGAAATGAGGATTAAGGAGATTC 1220  
 QY 401 HisGluAlaValGlyGlnIleMetSerLeuSerAlaIleThrProLysHisLeuLysSer 420  
 |||||||  
 Db 1221 CATGAGCTGTTGGGAAATCATGTCTCTCTGACGACACCTTAAGCATTTAAATTC 1280  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440  
 |||||||  
 Db 1281 ATTGCTCTCTGTCACCCGATTTTCAAGAAACAAATGAAACAGAAATTAACCTCTCTC 1340  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysThrArg 460  
 |||||||  
 Db 1341 AAACAGACACTCCAGATTGTTGGGACTCTGCCATTACTTACATGTTAGAGAAATGTGAGG 1400  
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpGluMet 480  
 |||||||  
 Db 1401 TGGATGCTCTTTAAAGGGGAATTCGCCAAAGACAGCTGATGAAAGTGTGGGAGATG 1460  
 QY 481 LysArgGluIleValGlyValGluProValProHisAspGluThrTyrCysAspPro 500  
 |||||||  
 Db 1461 AAGCAGAGATAGTTGGGGTGTGGAGCTGTGCCCATGATGAACACTACTGTGACCCC 1520  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520  
 |||||||  
 Db 1521 GCATCTCTGTCATGTTCTTCAATGATATCTCATTTCCATTTACACAGAACACCTT 1580  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlnGlnAlaIleAlaLysHisGluGlyProLeuHis 540  
 |||||||  
 Db 1581 TACCAATTCAGATTTCAGAGAAGCACTTGTCAAGACAGCTAAACATGAAGCCCTCTGAC 1640  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560  
 |||||||  
 Db 1641 AAATGTGACATCTCAACACTCAAGAACCTGACAGAACCTGCAATATCTGAGGCTT 1700  
 QY 561 GlyLysSerGluProTyrThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsn 580  
 |||||||  
 Db 1701 GGAAATATAGAAACCTGGAACCTAGACATTGGAAATGTTGTAGAGCAAAAGACATGAT 1760  
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGlnAsnLys 600  
 |||||||  
 Db 1761 GTAAGCCACTGCTCAACTACTTGTAGCCCTTATTTACCTGGCGAAGACCAAGCAAG 1820  
 QY 601 AsnSerPheValGlyTyrSerThrAspThrSerProTyrAlaAspGlnSerIleLysVal 620  
 |||||||  
 Db 1821 AATCTTTTGTGGATGAGTACGACCTGAGTCCATATGCAACCAAGCATCAAGTG 1880  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyLysPylsAlaTyrGluTyrPasnAspAsnGluMet 640  
 |||||||  
 Db 1881 AGGATTAAGCCTAAATTCAGCTCTCTGTGAGATTAAGCATATGATGAAGCAACATGAATG 1940  
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660  
 |||||||  
 Db 1941 TACCTGTTCCGATCATCTGTTGACATATCTATGAGGAGTACTTTTAAAGTAAATAAT 2000  
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 |||||||  
 Db 2001 CAGATGATTTCTTTTGGGAGAGAGATGTCCAGTGGCTAATTTGAACCAAGATCTCC 2060  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700  
 |||||||  
 Db 2061 TTTAATTTCTTTGTACAGCAGCACTAAATATGTCTGATATCATCTCAGAACTGAAGT 2120  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 |||||||  
 Db 2121 GAAAAGGCCATCAGATGTCCCGGAGCCGTATCAATGATGCTTCCGTCTGATAGACAAAC 2180  
 QY 721 SerLeuGluPheLeuGlyLysIleProThrLeuGlyProProAsnGlnProProValSer 740  
 |||||||  
 Db 2181 ACCCTAGAGATTTCTGGGGATACAGCAACACTTGGACCTCTTAACCAAGCCCTCTGTTCC 2240

QY 741 IleTrpLeuIleValPheGlyValIleMetGlyValIleValIleValIleLeu 760  
 |||||||  
 Db 2241 ATATGCTGATTTCTTTTGGAGTTGTGATGGAGATCATGTGTGGCATTTCTCATCTCG 2300  
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro 780  
 |||||||  
 Db 2301 ATCTTCAGTGGATCAGAGATCGAAGAGAAAGAAATTAAGCAAGATGAGAGAAATCTC 2360  
 QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAspProGlyPheGlnAsnThrAspAsp 800  
 |||||||  
 Db 2361 TATGCTTCATCGATATTTGCAAGAGAAATTAATCCAGATTCCTCAAAACACTGATGAT 2420  
 QY 801 ValGlnThrSerPhe 805  
 |||||||  
 Db 2421 GTTCAGACCTCTCTT 2435  
 RESULT 7  
 AX047758  
 LOCUS AX047758 3334 bp DNA linear PAT 15-DEC-2000  
 DEFINITION Sequence 1 from Patent WO0070032.  
 ACCESSION AX047758  
 VERSION AX047758.1 GI:11876765  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 3334)  
 AUTHORS Piddington, C.S., Petric, C.R., Shoemaker, K.E. and Bishop, P.D.  
 TITLE Zacez: a human metalloenzyme  
 JOURNAL Patent: WO 0070032-A 1 23-NOV-2000;  
 ZymoGenetics, Inc. (US)  
 FEATURES  
 source location/Qualifiers  
 1..3334  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 35..2452  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC18991.1"  
 /db\_xref="GI:11876765"  
 /translation="MSSSWLLSLVAVTAOSTIEBQAKFFLDKFNHEDLPYQSS  
 LSWNTNTIENVOQNMNAGDKWSEAELEQSTIAQNTPIQETQNTLVKIQALQLOO  
 NCSYVLEDKSRRLNTIINTMSTIYSTGKVCMPDQCLLEPGLNIMANSIDLYNE  
 RLWAMESRSEVQKOLRPLYEEYVVLKEMKAMANYEDYGYMRDYEIVNVDYDS  
 RGQLEDVHTETEEELKPLYEHLHAYVRKAKMAVYSYISPGCLPAHLIGMGREFMT  
 NLSLTPRGOKPNDIVDAMVDOAMDORLPEKEAEKFEFVSGLNOMGFEWNSMLT  
 DGRNOKAVCHPTAMDIGKGRIRILMCKRYMTPDLTAHHEMGIQYMAVAAQPELL  
 KGNANGKEHAEGLMSLSAATPKHLKSLGLSPDQEDNETELNFKLQALTYGTL  
 PPTYLEKRRMVFEGEIPKDWMRKWMKREIVGVEPVPHDETCDPASPILRSVSN  
 DYSEIYRTLYOFQFOALCOAAKHGEPLEKCDISNTEAGOKLFNMLRIGSEPM  
 TLALENVGAKNMNVRLNRYPEFTWLDQNKSPFGWSDMSPYVDQSIKVISL  
 KSLGDKAVEMNDENYLFERSVAVARQYFLKYKNOMILFEEDVRAVANKPRISFN  
 PFTYAKNYSDIIPREYVKAIRMSRRKINDAFRINDSLEFLGIQPLGSPNOPVS  
 IMLIVGCVMGVIVGIVLIFTGTRDKRKKKAKASGENPASYDISGNNPGRQNT  
 DDVQISF"  
 BASE COUNT 1011 a 640 c 754 g 929 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 3334  
 Score: 4291.00 Matches: 805  
 Percent Similarity: 100.00% Conserved: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-978-385-2 (1-805) x AX047758 (1-3334)  
 QY 1 MetSerSerSerThrLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20  
 |||||||  
 Db 35 ATGTCAGAGCTTCTTCTGCTCTTCTGAGCCTTGTCTGCTGTAACCTGCTGCTGACGACC 94



QY 21 IleGluGluGlnAlaIleAspThrPheLeuAspLysPheAsnHisGluAlaLeuAspLeuPhe 40  
 |||||  
 Db 95 ATTAGAGAACAGCCCAAGACATTTTGGACAACTTTTAACCCAGAACCCGACCTGTTTC 154  
 QY 41 TyrGlnSerSerIleuAlaSerTyrPasnTyrAsnThrAsnIleThrGluGluAsnValGln 60  
 |||||  
 Db 155 TATCAAGATTCACCTTCCTTCTTGGAATATTAACACCAATATTAATCTGAAAGAAATGTCCAA 214  
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuIleGluGlnSerThrLeuAla 80  
 |||||  
 Db 215 AACATGAATATCTCTGGGACCAATGCTGCTTTTAAAGAACAGTCCACACTTGGC 274  
 QY 81 GluMetLysProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnIleuGlnAlaLeu 100  
 |||||  
 Db 275 CAATGTATTCACCTACAGAAATTCAGAAATTCACAGTCAAGCTTCAGCTGCAGCTCTT 334  
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120  
 |||||  
 Db 335 CAGCAAAATGGGTCTTCAGTGGCTCTCGAAGACAAAGCAAGCGTTGAAACAATTTCTA 394  
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140  
 |||||  
 Db 395 AATACAAATAGCACCCATCTACACTGAGTGAAGAAAGTTGTAAACCCAGATATCCACAAAGAA 454  
 QY 141 CysLeuLeuLeuGluIleProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160  
 |||||  
 Db 455 TGCTTATTAATCTGAACACAGTTTGAATGAATAATGCGCAACAGTTTAACTTACATGAG 514  
 QY 161 ArgLeuThrAlaIleTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180  
 |||||  
 Db 515 AGGCTCTGGGCTGGGAAAGCTGGAGATCTGAGGTGCGGCAACAGCTGAGCGCAATATAT 574  
 QY 181 GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200  
 |||||  
 Db 575 GAAGAGTATGTGGTCTTGAAATAGAGATGAGACGACAAATCATTTAGAGACTATGGG 634  
 QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220  
 |||||  
 Db 635 GATTAATGGAGAGAGACTATGACATAAATGGGGTAGATGCTATGACTACAGCGCGGC 694  
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240  
 |||||  
 Db 695 CAGTTGATTTGAAGATGTGGAACATCTTGAAGAGATTAACCAATTAATGACATCTT 754  
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260  
 |||||  
 Db 755 CATGCCATATGTAGGCAAGATTGATGAATGCCATCTCTCTATATACAGTCCAAATTGGA 814  
 QY 261 CysLeuProAlaHisLeuLeuGlyLysPheTrpGlyArgPheTrpThrAsnLeuTyrSer 280  
 |||||  
 Db 815 TCCCTCCCTGGCTCATTTGGTGTGATATGGGGGTAGATTTTGGCAAAATCTGTACTCT 874  
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 |||||  
 Db 875 TTGACAGTTTCCCTTTGGACAGAAACCAACATAGATGTACTGATGCAATGGGTGACCA 934  
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320  
 |||||  
 Db 935 GCCCTGGGATGACAGAGATATTTCAAGAGAGCGCGAGAGTCTTGTATCTGTGGTCTT 994  
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340  
 |||||  
 Db 995 CCTATATAGACCAAGCATTTCTGGGAAATTCATGCTTAACAGGACCCGAAATGTTCTAG 1054  
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 |||||  
 Db 1055 AAAGAGCTGTCCATCCACAGCTTGGGAGCTGGGAGAGGCGCACTTCAAGATCCTTAATG 1114  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisIleGluMetGlyHisIleGln 380  
 |||||  
 Db 1115 TGCACAAAGGTGACATGAGACTTCTGTGACAGCTCATCTAGATGGGCGCATATCCAG 1174  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400  
 |||||  
 Db 1175 TATGATATGGCATATGTGCAACACTTTTCTGTCAAGAAATGAGCATTAAGAGGATTC 1234  
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisIleLysSer 420  
 |||||  
 Db 1235 CATGAGCTGTGGGGAAATCATGTCACTTTCTGAGGCACTTAAGCATTTTAAATCC 1294  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440  
 |||||  
 Db 1295 ATTTGCTTCTGTCAACCCGATTTTCAAGAAACATGAACAGAAATTAACCTCTCTGCTC 1354  
 QY 441 LysGlnAlaLeuThrIleValGlyThrPheProPheThrTyrMetLeuGluLysTrpArg 460  
 |||||  
 Db 1355 AAACAGACACTACGATTTGTGGGACTGTGCCATTTACTTACATGATGAGAAATGTGAGG 1414  
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpGluMet 480  
 |||||  
 Db 1415 TGATGTCTTTTAAAGGGAAATTCACCAAGACAGTGGATGAAAAAGTGGTGGGAGATG 1474  
 QY 481 LysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCysAspPro 500  
 |||||  
 Db 1475 AAGCAGAGATAGTTGGGGGTGTGGAACTGTGCCCATGATGAACATACTGTGACCCC 1534  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520  
 |||||  
 Db 1535 GCATCTCTGTCATGTTTCTTAATGATTAATCAATTCATTCATGATTTTACAAAGAGCCCTT 1594  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaIleAsnHisGlnGluProLeuHis 540  
 |||||  
 Db 1595 TACCAATGTCAGTTTCAAGAAAGCACTTTGTCAACACACTGAACATGAAGGCCCTTCAC 1654  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560  
 |||||  
 Db 1655 AAATGTGACATCTCAACTCAACTTACAGAGCTGGACAGAAAGCTGTTCAATATGCTGAGCCTT 1714  
 QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGluLysAsnValValGlyAlaLysAsnMetAsn 580  
 |||||  
 Db 1715 GGAAATACAGAACCTGAGACCTGACCTGACATTTGGAATAATTTGTAGAGCAAAAGCAATGAA 1774  
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGlnAsnLys 600  
 |||||  
 Db 1775 GTAAGGCCACTGCTCAACTTACTTTGAGCCCTTATTTACTGTGGCGAAGACCAAGCAAG 1834  
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620  
 |||||  
 Db 1835 AATTTCTTTTGTGGATGAGAGTACCGACTGGAGTGCATATGACAGCAAAACATCAAGTG 1894  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640  
 |||||  
 Db 1895 AGATTAAGCCTAAATATCAGCTCTTGGAGATTAAGCATATGAAAGCAATGAAATG 1954  
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660  
 |||||  
 Db 1955 TACCTGTTCCGATCATCTGTTGATATGATGATGAGGACAGTCTTTTAAAGTAAAAAAT 2014  
 QY 661 GluMetIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 |||||  
 Db 2015 CAGATGATTTCTTTTGGGAGAGAGATGTCGAGTGGCTAATTTGAAACCAAGAAATCTGC 2074  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700  
 |||||  
 Db 2075 TTTAAATTTCTTTGCACTGCACTAAATAATGTCTGATATCAATTCCTGAACCTGAAGTT 2134  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 |||||  
 Db 2135 GAAMAGCCATCAGAGATGCTCCGAGCCGATATGAATGATCTTTCCGTCTGAATGCAAC 2194  
 QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 |||||  
 Db 2195 AGCTTAGATTTCTGGGAGATACAGCAACACTTGGACCTTCATACCAAGCCCTGTTTCC 2254  
 QY 741 IleTrpLeuIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeu 760  
 |||||



Db 2255 ATATGCGCTGATTTGTTTGGAGTGTGATGGAGTGAATGATGCTTGCATCTCCG 2314  
 Qy 761 TlephethGlylleArgaspArgLysLysAsnLysAlaArgSerGlyLysnPro 780  
 Db 2315 ATCTTCACAGGATCGAGATCGAAGAAATAAAGCAAGAGAGAAATATCT 2374  
 Qy 781 TyrAlaSerLleAspLleSerLysGlyLysAsnProGlyPheGlnAsnThrAsp 800  
 Db 2375 TATGCTCCATGATATTGACAAAGAGAAATATCCAGATTCCTCAAAACACTGATGAT 2434  
 Qy 801 ValGlnThrSerPhe 805  
 Db 2435 GTTCAGACCTCCTT 2449  
 RESULT 8  
 ARI35177  
 LOCUS ARI35177 3396 bp DNA linear PAT 16-MAY-2001  
 DEFINITION Sequence 1 from patent US 6194556.  
 ACCESSION ARI35177  
 VERSION ARI35177.1 GI:14124082  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 3396)  
 AUTHORS Action, S. Laurene, and Robison, K. Earl.  
 TITLE Angiotensin converting enzyme homolog and therapeutic and  
 diagnostic uses therefor  
 JOURNAL Patent: US 6194556-A 1 27-FEB-2001:  
 FEATURES  
 source Location/Qualifiers  
 1..3396  
 /organism="unknown"  
 BASE COUNT 1034 a 659 c 772 g 931 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 3396  
 Score: 4291.00 Matches: 805  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Db: Gaps: 0  
 US-09-978-385-2 (1-805) x ARI35177 (1-3396)  
 Qy 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20  
 Db 82 ATGTCAAGCTCTTCTGCGCTCTTCAGCTGTGCTGTAAGTCTGCTCACTCACTCACC 141  
 Qy 21 TleGlnGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaLysPhe 40  
 Db 142 ATTGAGAGACAGCCCAAGACATTTTGGACAACTTAAACACAGAACCCGAGACCTGTC 201  
 Qy 41 TyrGlnSerSerLeuAlaSerTrpAsnTrpAsnThrAsnLleThrGlnGlnLysnValGln 60  
 Db 202 TATCAAAAGTCACTGCTCTTCTTGAATTTATACCAATATTTACTGAAGAGATGTCCAA 261  
 Qy 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla 80  
 Db 262 AACCTGATATATGCTGGGACCAATGTGCTCTTTTAAAGAAACAGTCCACACTTGGC 321  
 Qy 81 GlnMetTrpProLeuGlnGlnLleGlnAsnLleThrValLysLeuGlnLeuGlnAlaLeu 100  
 Db 322 CAATATGATCCACTACAGAAATTCAGAAATTCACAGTCAAGCTTCAAGCTGCGCTCTT 381  
 Qy 101 GlnGlnAsnGlySerSerValLeuSerGlyAspLysSerLysArgLeuAsnThrLleLeu 120  
 Db 382 CAGCAAAATGAGTCTTCAAGCTGTGCTGAGAACAGAACAGAAAGCTTGAACACAAATTCYA 441  
 Qy 121 AsnThrMetSerThrLleTrpSerThrGlnLysValCysAsnProAspAsnProGlnGln 140  
 Db 442 AATATCAATGAGACACATCTACAGTACTGGAAAGTTTGTAACCAAGATTAATCCACAGAA 501

Qy 141 CysLeuLeuLeuGlnProGlyLysLeuAsnGlnLleMetAlaAsnSerLeuAspTrpAsnGln 160  
 Db 502 TGCTTATTACTTGAACACAGGTTTGAATGAATATATGCAAAACAGTTTACACTCAATGAG 561  
 Qy 161 ArgLeuTrpAlaTrpGlySerTrpArgSerGlyValGlyLysGlnLeuArgProLeuTrp 180  
 Db 562 AGGCTCGGCGCTGGGAAGCTGGAGATCTGAGCTCGGCAAGACAGCTGAGAGCCATTATAT 621  
 Qy 181 GlnGlnTrpValValLleLeuLysAsnGlnMetAlaArgAlaAsnHisTrpGlyLysAspTrp 200  
 Db 622 GAGAGATATGTGCTTGAATAATGACATGACATGCTTGAAGAGATTAACCATTAATGAAACATCTT 801  
 Qy 201 AspTrpTrpArgGlyAspTrpGlyValAlaAsnGlyValAspGlyTrpAspTrpSerArgGly 220  
 Db 682 GATTATTGAGAGAGAGACTATGAGTAAATGGGGTAGATGGCTATGACTACAGCCCGGC 741  
 Qy 221 GlnLeuLleGlyAspValGlnHisThrPheGlnGlnLleLysProLeuTrpGlnHisLeu 240  
 Db 742 CAGTTGATTTGAAGATGTGGAAACATACCTTGAAGAGATTAACCATTAATGAAACATCTT 801  
 Qy 241 HisAlaTrpValArgAlaLysLeuMetAsnAlaTrpProSerTrpLleSerProLleGly 260  
 Db 802 CATGCTTATGTGAGGCAAGATTCATGAAATGCTTCTTCTATATCACTCAATGGA 861  
 Qy 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyValArgPheTrpThrAsnLeuTrpSer 280  
 Db 862 TGCCCTCCCTGCTCATTTCTTCTGATATGAGGGATGAGATTTGGCAAAATCTGTACTCT 921  
 Qy 281 LeuThrValProPheGlyGlnLysProAsnLleAspValThrAspAlaMetValAspGln 300  
 Db 922 TTGACAGTCCCTTTGGACAAACCAACCAATCAATGTCTGTGATGCAAAAGGTGGACAG 981  
 Qy 301 AlaTrpAspAlaGlnArgLlePheLysGlnAlaLysLysPhePheValSerValGlyLeu 320  
 Db 982 GCTTGAGATGACAGAGAAATTCAGAGAGCCAGAAAGTTCTTGTATCTGTGCTT 1041  
 Qy 321 ProAsnMetThrGlnGlyPheTrpGlyLysnSerMetLeuThrAspProGlyAsnValGln 340  
 Db 1042 CCTAATATGACTCAAGGATCTCTGGGAAATTCATGTCTAAAGGACCCAGCAAAATGTTTCAG 1101  
 Qy 341 LysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgLleLeuMet 360  
 Db 1102 AAAGCAGTCTGCATCCACAGCTTGGGACCTGGGAGAGGCGACTTCAGAGATCTTATG 1161  
 Qy 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHisLleGln 380  
 Db 1162 TGCAAAAGAGTGAACATGAGAGACTCTCTGACACTGATCATGATGAGGCAATATCCAG 1221  
 Qy 381 TyrAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyLysPhe 400  
 Db 1222 TATGATATGCAATATGCTGCACAAACCTTTCTGTAGAAATGAGCTAATGAGAGATTC 1281  
 Qy 401 HisGlnAlaValGlyGlnLleMetSerLeuSerLleAlaTrpProLysHisLeuLysSer 420  
 Db 1282 CATGAGAGCTGTGGGAAATATGATGCTTCTGAGCCACACCTTAAGCATTTTAAATTC 1341  
 Qy 421 IleGlyLeuLeuSerProAspPheGlnGlyLysAsnGlnTrpGlnLleAsnPheLeu 440  
 Db 1342 ATTGCTTCTGTCACCCGATTTTCAAGAACATGAAACAGAAATTAACCTCTGCTC 1401  
 Qy 441 LysGlnAlaLeuThrLleValGlyThrLeuProPheThrTrpMetLeuGlnLysTrpArg 460  
 Db 1402 AAACAGACACTCACATTTGTTGGGCTCTGCAATTTACTTTCATGCTTGAAGAAATGGAGG 1461  
 Qy 461 TrpMetValPheLysGlyGlnLleProLysAspGlnTrpMetLysLysTrpTrpGlnMet 480  
 Db 1462 TGGATGGCTTTTAAAGGGAATTTCCAAACACAGTGAATGAAAGTGGTGGAGATG 1521  
 Qy 481 LysArgGlnLleValGlyValAlaGlnProValProHisAspGlnThrTrpCysAspPro 500  
 Db 1522 AAGCAGAGATATGTTGGGGTGGTGAACCTGCCCCATGATGAACATACTGAGACCC 1581  
 Qy 501 AlaSerLeuPheHisValSerAsnAspTrpSerPheLleArgTrpThrArgThrLeu 520



Db 502 TGGTATATCTTGAAACGAGTTTGAATGCAATATGCGCAACGATTGACTACAAATGAG 561  
 QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyGlnLeuArgProLeuTyr 180  
 Db 562 AGGCTCTGGGCTTGGGAAAGCGGAGATCTGAGGCTGGCAACGACCTGAGGCAATTAAT 621  
 QY 181 GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200  
 Db 622 GAAAGCATATGCTGCTTGAATAATGAGATGCGAAGCAAAATCATTAAGAGACTATGAG 681  
 QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220  
 Db 682 GATTTATGAGAGAGAGACTATGAGAAATGAGGATAGAGCTATGAGCTACAGCCGCGC 741  
 QY 221 GlnLeuIleGluAspValGlnHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240  
 Db 742 CAGTTGATTGAAGATGTGGAACATCTTTGAAAGATTTAAACCATTTATGAAACATGTT 801  
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260  
 Db 802 CATGCCATATGAGGCGCAAAAGTTGATGAATGCCCTATCCTTATATCATGCTCAATGGA 861  
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280  
 Db 862 TGGCTCCCTGCTCATTTCTGCTGATATGTGGGATGATTTTGGACAAATCTGTACTCT 921  
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 Db 922 TTGCAGCTTCCCTTTGGACAGAAACCAATGATGTTACTGATGCAATGGTGGACAG 981  
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGlnLysPhePheValSerValIleLeu 320  
 Db 982 GCGTGGGATGCAACAGAAATATTCAGAGGCGCGAAGCTTTCTGTATCTGTGCTCT 1041  
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340  
 Db 1042 CCTAATATGATCAACAGATTTCTGGGAAATTCATGCTAACGACCGCAAGAAATCTTCA 1101  
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 Db 1102 AAGAGCTGCTGCCATCCCAAGCTTGGAGCTGGGGAAGGCGACTTCAGATCTTATG 1161  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380  
 Db 1162 TGGACAAAGGTGACATGACGACGCTCTGACAGCTCATCATGAGATGGGCAATACAG 1221  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyValAsnGluLysPhe 400  
 Db 1222 TATGATATGCGATATGCTGCACAACTTTCTGCTAAGAAATGAGCTAATGAAAGATTC 1281  
 QY 401 HisGluAlaValGlyLysIleLeuMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
 Db 1282 CATGAAGCTGTGGGAAATTCATGCTCTTCTGCAGCGCAACCTTAAGATTTAAAAATCC 1341  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440  
 Db 1342 ATGGCTCTCTGCAACCCGATTTTCAAGAAAGCAATGAAGAAATAAATCTCTGCTC 1401  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTrpArg 460  
 Db 1402 AAACAAGCACTACGATTTGGGACTGTGCAATTTACTTAATGATGAGAAAGTGGAGG 1461  
 QY 461 TrpMetValPheLysGlyLysIleLeuProLysAspGlnTrpMetLysLysTrpArgGluMet 480  
 Db 1462 TGGATGCTCTTAAAGGGAATTCACAAAGACAGCTGATGAAAAAGTGGGAGATG 1521  
 QY 481 LysArgGluIleValGlyValGluProValProHisAspGluThrTyrCysAspPro 500  
 Db 1522 AAGCGAGAGATAGTGGGCTGGGAACTGTGCCCAAGATGAAGAAACATACGTGACCCC 1581  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520  
 Db 1582 GCATCTCTGCTCATGTTCTTAATGATTAATCATTCATTCATTCATTCATTCATTCATTCAT 1641

QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHis 540  
 Db 1642 TACCATATTCAGATTTTCAAGAAAGCACTTTGCAAGCACTTAAACATGAGAGCCCTTGCCAC 1701  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560  
 Db 1702 AAATGTGACATCTCAAACTGTACAGAGAGCTGAGACAGAAACTGTTCAATATGCTGAGGCTT 1761  
 QY 561 GlyLysSerGluProTyrPheLeuAlaLeuGluLysAsnValValGlyAlaLysAsnMetAsn 580  
 Db 1762 GGAATAACAGAACTCTGAGCCCTAGCATTTGAAATGTTGTAGACCAAGAACATGAAAT 1821  
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600  
 Db 1822 GTAAGCCACTGCTCAACTACTTTGAGCCCTTATTTTCTGCTGCTGGAAGACCGAAGAAC 1881  
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620  
 Db 1882 AATTCCTTTTGGGATGAGAGACCGAGTGGAGTCCATATGACAGCAACCATCAAAATG 1941  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640  
 Db 1942 AGGATTAAGCCCTAAATGAGCTCTTGGAGATTAAGCATATGAAATGGAACGACATGAAATG 2001  
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValIleAsn 660  
 Db 2002 TACCTGTCCGATATCTGTTGCTATGCTATGAGGAGATGCTTTTAAAAAGTAAAAAT 2061  
 QY 661 GlnMetIleLeuPheGlnGlyGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 Db 2062 CAGATGATTTCTTTTGGGAGAGAGATGTGCGAGTGGCTAATTTGAAACCAAAATCTCC 2121  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700  
 Db 2122 TTTAATTTCTTTGCTGCTGACCTTAATAATGTGCTGATTCATCTCTTAACACTGAATG 2181  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 Db 2182 GAAGAAGGCATCAGATGTCGCCGAGCGATATCAATGATCTTCGCTCGAATGACAAAC 2241  
 QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 Db 2242 AGCCTAGATTTCTGGGGAATACACCAACTGGACCTCTTAACCGCCCTGTTTCC 2301  
 QY 741 IleTrpLeuIleValPheGlyValIleMetGlyValIleValGlyIleValIleLeu 760  
 Db 2302 ATATGGCTGATGTTTGTGGAGTTGTATGGAGTGTATGTTGGTATGTCATTCCTG 2361  
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780  
 Db 2362 ATCTTCACTGGGATCAGAGATCGGAGAAAGAAATTAACAGCAAGATGGAGAAATCTT 2421  
 QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnTrpAspAsp 800  
 Db 2422 TATGCTCTCACTGATATTAAGCAAGAGAAATTAATCCAGATTTCCAAAACCTGATGAT 2481  
 QY 801 ValGlnThrSerPhe 805  
 Db 2482 GTTCAGACCTCTTT 2496  
 RESULT 10  
 AX431513 3396 bp DNA linear PAT 28-JUN-2002  
 LOCUS AX431513  
 DEFINITION Sequence 1 from Patent WO0239997.  
 ACCESSION AX431513  
 VERSION AX431513.1 GI:21656345  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1

**AUTHORS** Gould, A. E., Stricker-Krongrad, A., Acton, S. L., Brown, J. A., Guan, B., Dales, N. A., Kadamby, V. J., Ocain, T. D., Patane, M. and Solomon, M.

**TITLE** Ace-2 modulating compounds and methods of use thereof

**JOURNAL** Patent: WO 023997-A 1 23-MAY-2002;

FEATURES	Location/Qualifiers
source	1. .3396

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
82. .2499
CDS

```

/protein\_id="CAD37288.1"  
 /db\_xref="GI:21656346"  
 /translation="MSSSSWLLSLVAVTAQSTIEQAKTFILDKENHEADLFYQSS  
 LASNVTNITNEEVOVMNNMAGDKSAEFLKQSTQSPVLAQELQNTVTKLQALQALQ  
 NGSSVLSEDKSKRLNTLTNTSTSTSTGKGVQNPNNPQECILLRGLNEMANSLDYNE  
 RLAMESWRSEVGKQLPRIVEEYVVLNEMARAHYEDGYRMGDEYVANGSDTSDS  
 RGQILDEYEHFEEIKPLVEHLHAYAKLMAARLTSTISPLGLPAHLHLCGMRFWPT  
 NLVSLTFLPGCKPRNIDYTDKAVDQAMQARLFEKAEKFEVSGLNPNTQGMENSMULT  
 DGVNQKAVCHPTAMDQKGDGFRILMCTQYKIMDFDLAHNHMGHIVDYMVAAPQFL  
 RGNBEGFHEAVGDIEMSLSATPKHLSIGLSEDFQEDNETINFILKALTLTGITIT  
 PFTMLERKMRWYKGEIPKQDMKMKWEMKREIVGVEYEPHEQTCDDPASHFNHS  
 DYSTIRYRTLYQFOEALCOAKHHEGFLHKDLSINTEPAGQGLNNLTIRKSEPM  
 TLALENVGAKNNMVRPLNTEPLTFMLDQONKNSPYGSDMSPAQDSIKRITST  
 KSAAGDKAEYNNDEMYLFRSSVAYARQFLAKNNQMLGEDDYVAANIKPISFN  
 FEYVAPKPNVDIIRTEYKAKIRMSRSINDAFPLNNLSLEFLDIQPLVLPNPNPVS  
 IMLVEYGVKVIIVGLIILLFTGIRDKKKNNKRSNGENPYASIDISKENNPEQNT  
 DVDQVTSF

Alignment Scores:	
Pred. No.:	0
Score:	4291.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
Length:	3396
Matches:	805
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-978-385-2 (1-805) x AX431513 (1-3396)

QY	1	MeSerSerSerSerTripleLeuLeuLeuSerLeuValAlaValThrAlaIaInSerThr	20
Db	82	ATGTCAAGCTCTTCTCGGCTCCTCTTCAGCCTGTGGCTGTACTCTGCTCAGTCACC	141
QY	21	IIeGIuGIuGIIaIeLySThrPheLeuAspLysPheAsnHISgluIIaGluAspLeuPhe	40
Db	142	ATTGAGAACACAGCCAAACACATTTTGGACAGCTTTAACCCAGAACCCGAACCTGTTC	203
QY	41	TYrGInSerSerLeuAlaSerTyrAsnTyrAsnThrAsnIleThrGIuGluAsnValGIn	60
Db	202	TATCAAAAGTTCCTCTGCTTGGAAATTATAACACCACATATTACGAGAGAGATGTCCAA	263
QY	61	AsnMetAsnAsnAlaGIIaAspLysTyrSerAlaPheLeuLysGIuGInSerThrLeuAla	80
Db	262	AACATGATATATGCTGGGGGACAAATGGTGTGCTTTTAAAGAACACGTCCACTTGCC	321
QY	81	GIuMetTyrProLeuGInGIIaIeGIaAsnLeuThrValLysLeuGIIeGInGIIaLeu	100
Db	322	CAAAATGATACCTACACAGAAATTCACAAATCTACAGTCCAGCTTCAGCTCAGGCTTT	381
QY	101	GIuGIIaAsnGIIySerSerValLeuSerGIuAspLysSerLysArgLeuAsnThrIleLeu	120
Db	382	CAGCAAAATGGGTCTCAGTCTGCTGCAGAAACAAAGCAAAAGCGTTGAACAACAATCTTA	441
QY	121	AsnThrMetSerThrIleTyrSerThrGIIyLysValCysAsnProAspAsnProGInGIIu	140
Db	442	AATACATAGAGACACATCTACAGTACGTGAAAGTTTGTATCCCGAATATCCCAAGAA	501
QY	141	CysLeuLeuLeuGIIuProGIIyLeuAsnGIIaIeMetAlaAsnSerLeuAspTyrAsnGIIu	160
Db	502	TGCTATTACTTGAACACAGGTTTGAAATGAAATATATGGCAACAGTTTGTACTACAAATGAG	561

QY	161	ATGLeuTPRAlaTPRcIuSeTTPRagSeGIuValGILySGInLeuARProLeuTYR	180
Db	562	AGGCTCTGGGCGTTGGGAAGCTGGAGATCTGAGGTCGGCAAGCACTGGAGCCATTATAT	621
QY	181	GIUGLUYRValValLeuLeuYASnGILmeTAlaRgAlaSnHISYRIGLusPlyRGLY	200
Db	622	GAAGAGATAGTGGCTTGTAAATAAGAGATGGCAAGCAACATATTATAGACATATGGG	681
QY	201	ASPTTYRTPRAGLYASPTYRGIuValASnGILYAlaSPGLYTYRASPTYSERAGLY	220
Db	682	GATATTGGAGAGGAGCATTAAGAGTAAATGGGGTAGATGGCATACTACACGCCGGCG	741
QY	221	GINLeuILeGLuASVAlGInHISThPhaGILuGInILeYSPROLeuTYRGIuHISLeu	240
Db	742	CAGTGTATTGAAGATGGCAACATACCTTTGAAGACATTTAACCATTTATATAAATCTT	801
QY	241	HISAlaTYRValARgAlaLYSLeuMeTASnAlaTYRProSeTYRILeSERProILeGLY	260
Db	802	CATGGCTATGTGAGGGCAAAATGATGAAGCTTATCCCTCATATACGTGCATATGGA	861
QY	261	CYSLeuPROAlaHISLeuLeuGILYASMeTTRGILARgPhetRPThASnLeuTYSER	280
Db	862	TGCCTCCCTGCTCATTTGCTTGGGATATGGGGTAGATTTTGGCAAAATCTGCACCT	921
QY	281	LeuThRValProPhaGILyGInLYSPROSnILeASVAlThRASPAlaMeTAlaSPGLN	300
Db	922	TTGACAGTTCCTTTGGACAAACCACATAGATGTTACGATGACATGGTGGACACG	981
QY	301	AlATPRASPAlaGInARgILIEPhelYSGILuAlGILYSPhePheValSERValGILYLeu	320
Db	982	GCTTGAGATGCACGAGAATATTCACAGAGCGGCAAAATCTTTGTATCTGTGGTCTT	1041
QY	321	PROASnMeTThRgInGILYPhetRPGLuASnSeMeTLeuThRASPProGILYASnValGIn	340
Db	1042	CTATATGACTCAAGGATTCGGGAATAATCCATGCTTACGGACCCAGGAATGTTTCAG	1101
QY	341	LYSAlaValCYSHISPRoThRAlATPRASPLEuGILYSGLYASPPheARgILEuMeT	360
Db	1102	AAAGAGCTGTCGCATCCCAACAGCTTGGGACCTGGGGAAAGGGCGACTTCAGATCCCTAATG	1161
QY	361	CYSThRValYRmeTASPRASPhelLeuThRALaHISHSIGLmeGILHISILeGLIn	380
Db	1162	TGCAACAAGGACAAATGGACGACTTCCTGACAGCTCATAGACATGGGCGCATATCCAG	1221
QY	381	TYRASPMetalARYAlaAlaGInProPhelLeuLeuARgASnGILYAlaSnGILYPha	400
Db	1222	TATGATATGGCATATGCTGCACAACCTTTTCTGTATAGAAATGGAGCTAATGAAGATTC	1281
QY	401	HISGILuValaGILYGIuILeMeTSERLeuSERAlaAlaThRPROLYSHISLeuLYSER	420
Db	1282	CATGAAGCTGTGGGAAATCATGTCACTTCTGGAGCCACACCTTAAACATTTAAATTC	1341
QY	421	ILeGLYLeuLeuSERPROASPRheGInGILuASpASnGILThGILYLeASPhelLeuLeu	440
Db	1342	ATTGGCTCTTCGACACCCGATTTTCAAGAAGACATGAAGAAGCAATTAATTCCTCTC	1401
QY	441	LYSILuAlaLeuThRILeValGILYThRLeuPROPhetThRYRmeTLeuGILYSTPRARg	460
Db	1402	AAACAACACCTCACGATTTGGGACTTCGCCATTACTTACATGTTAGAGAAATGGAGG	1461
QY	461	TRPmeTAlaPhelYSGILYILeRPOLYASPRgInTRPMeTLYSYSTTPRGLYmeT	480
Db	1462	TGATATGCTCTTAAAGGGGAATTCCCAAGACCGATGATGAAAAAGTGTGGGAATG	1521
QY	481	LYSARGLuILeValGILYValValGILuPROAlaPROHISASPRgInThRYCYASPRPro	500
Db	1522	AAGGACAGATAGCTTGGGCTGGGCAACCTGTGCCCATGATGAACAATACGTGACCCG	1581
QY	501	AlaSERLeuPhelHISValSERASnSPYRSeRPhelLeuARgTYRThRARgInThRLeu	520
Db	1582	GCATCTGTGTCATGTTTCTATAGATATTCATTCATTCGATTCGATATTCACAAAGACCTTT	1641



Db 344 CAATGTCCTCCACTACAGAAATTCAGAAATCTCAGACTCAAGCTTCAGTGCAGGGCTT 403  
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerIysArgLeuAsnThrIleLeu 120  
 Db 404 CAGCAAAATGGGTCTTCAGTCTCTCTCAGAAAGACAAAGGCAAGCGTTGAACACAAATTCGA 463  
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140  
 Db 464 AATCAAAATGAGCACCACATCTACAGTACGAGAAAGTTGTATACCCAGATATCCACAGAA 523  
 QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160  
 Db 524 TGCTATTACTGAAACCAAGTTTGAATGAATATGCAAAACGTTTAACTACATCAATGAG 583  
 QY 161 ArgLeuThrPalaThrGluSerTyrPargSerGluValGlyLysGlnLeuArgProLeuTyr 180  
 Db 584 AGGCTCGGGCTGGGAAAGCTGAGAGTCTGAGGTGGCAGACAGCTGAGGCCATTATAT 643  
 QY 181 GluGluTyrValValleuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200  
 Db 644 GAAAGATATGCTGCTTGAAGAAATGAGATGCGAAGAGCAAAATCATTTATGAGCATATGG 703  
 QY 201 AspTyrThrArgLysAspTyrGluValAsnGlyValAsnGlyTyrAspTyrSerArgLys 220  
 Db 704 GATTATGAGAGAGAGACTATGAGTAAATGGGGTAGATGGCTATGACTACACCGCGGC 763  
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240  
 Db 764 CAGTTATGTAAGATGTGGAACATCTTGAAGAGATTTAAACCATTTATGACATCTT 823  
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260  
 Db 824 CATGCTATGTAGGGCAAGTTGATGATGCTTATCTTCTATATGACCAATTTGA 883  
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTyrThrAsnLeuTyrSer 280  
 Db 884 TGCTCTCCCTCATTTGCTGTGATATGTTGGGGTAGATTTTGGCAAAATCTTACTCTT 943  
 QY 281 LeuThrValProPheGlyGluLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 Db 944 TTGACAGCTTCCCTTTGGACAGAAACCAACATAGTCTTACTGATGCAATGGTGGACAG 1003  
 QY 301 AlaTyrPaspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320  
 Db 1004 GCCGTGGATGACACAGATATTCAGAGAGGCCAGAAAGTTCTTGTATCTGTTGGCTT 1063  
 QY 321 ProAsnMetThrGlnGlyPheTyrPgluAsnSerMetLeuThrAspProGlyAsnValGln 340  
 Db 1064 CCTAATATGACTCAAGGATCTGGGAAATTCATGCTAACGGACCCAGAAATGTTTCAG 1123  
 QY 341 LysAlaValCysHisProThrAlaTyrPaspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 Db 1124 AAAGCAGCTCCCATCCACAGCTTGGGACCTGGGGAAGGGCCAGCTTCAGAGATCTTATG 1183  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlnGluMetGlyHisIleGln 380  
 Db 1184 TGCAACAAGGGAGCAATGAGAGACTCTCTGACAGCTCATCATAGATGGGGCCATATCCAG 1243  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400  
 Db 1244 TATGATATGGCATATGCTGCAACAACCTTCTCTAAGAAATGAGACCTAATGAGGATTC 1303  
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
 Db 1304 CATAAACCTGTGGGAAATCATGTCATTTTGCAGCCACACACCTAAGCATTTAAAAATCC 1363  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAsnGlyThrGluIleAsnPheLeuLeu 440  
 Db 1364 ATTAGGTCTTGTACCCCGATTTTCAAGAGCAAAATGAAGAAATAACTTCTCTGCTC 1423  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTyrArg 460  
 Db 1424 AAACAAGCACTCAGATTTGTTGGACTGTGCCATTTACTTACATGTGTAGAGAGTGGAGG 1483

---

QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTyrTrpGluMet 480  
 Db 1484 TGATGTGCTCTTAAAGGGAAATTTCCCAAGACCAGTGGATGTAAGAAAGTGGGGAGATG 1543  
 QY 481 LysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCysAspPro 500  
 Db 1544 AAGGAGAGATAGTGGGGTGGTGGAAACCTGTCCCATGATGAAACATACACTGTGACCC 1603  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520  
 Db 1604 GCATCTCTGTCCATGTTTCTTAATGATTTACTCATTCATTCGATATACAGAGACCCCTT 1663  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis 540  
 Db 1664 TACCAATTCAGATTTCAGAGAGCACTTGTTCAGAGCAGCTTAAACATGAGGCCCTTGGAC 1723  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560  
 Db 1724 AAATGTGACATCTCAAACTCTACAGAGCTGGACAGAAACCTGTCAATATGCTGAGGCTT 1783  
 QY 561 GlyLysSerGluProTyrThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsn 580  
 Db 1784 GGAATTCAGAACCTCGAACCTGACATTCGATGGAATGTTGTATGAGCAAGAAACATGAT 1843  
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGlnAsnLys 600  
 Db 1844 GTAAGCCACCTGCTCAACTACTTATGAGCCCTTATTTACTGGCTGTAAGACCAACAG 1903  
 QY 601 AsnSerPheValGlyTyrSerThrAspTyrPargSerProTyrAlaAspGlnSerIleLysVal 620  
 Db 1904 AATCTTTTGTGGATGAGTACCGACCTGAGTCCATATGCAACCAAAAGCATCAAAAGTG 1963  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTyrPasnAspAsnGluMet 640  
 Db 1964 AGGATTAAGCCTAAATTCAGACTCTTGAGATTAAGCATATGATGATGAACACGAAATGAATG 2023  
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLysValLysAsn 660  
 Db 2024 TACCTGTTCGATCATTCTGTCATATGCTATGAGGACAGTCTTTTAAAAAGTAAAAAT 2083  
 QY 661 GlnMetIleLeuPheGlyGlyGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 Db 2084 CAGATGATCTTTTGGGAGAGAGATGTGGCACTGGCTAATTTGAAACCAAGAAATCTCC 2143  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700  
 Db 2144 TTATATTTCTTGTACAGCTGACCTTAAATATGTGTCTGATCATCTCTAGAACTGAAGTT 2203  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 Db 2204 GAAAAGGCATCAGAGATGTCCCGGAGCGGTATCAATGATGCTTCCGTGTAATGCAAC 2263  
 QY 721 SerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 Db 2264 AGCTTACAGTTTCTGGGATACAGCCAACTTGAGACTTCCTACACAGCCCTCTGTTTC 2323  
 QY 741 IleThrPheLeuIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeu 760  
 Db 2324 ATATGGCTGATGTGTTTGGAGTTGTGATGGAGTATGTGTGGCTATGTGATCTCAG 2383  
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysLysPro 780  
 Db 2384 ATCTTCACTGGGATCAGAGATCGGAAGAAAGAAATTAAGCAAGAAAGTGGAAATATCCT 2443  
 QY 781 TyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAsp 800  
 Db 2444 TATGCTCTCATGTATTTGCAAGAGAAATATATCCAGATTTCCAAAACACTGATGAT 2503  
 QY 801 ValGlnThrSerPhe 805  
 Db 2504 GTTCAGACCTCTCTTT 2518



RESULT 12  
 AB046569  
 LOCUS  
 DEFINITION Homo sapiens ace2 mRNA, complete cds.  
 ACCESSION AB046569  
 VERSION AB046569.1 GI:13516971  
 KEYWORDS  
 SOURCE Homo sapiens adult cDNA to mRNA, clone\_11b:Sugano human cDNA  
 ORGANISM  
 Homo sapiens  
 Library Kala clone:kala4505.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 1 (sites)  
 2 (bases 1 to 2599)  
 AUTHORS Suzuki, Y., Matanabe, M. and Sugano, S.  
 TITLE Cloning, expression analysis and chromosomal localization of a novel ACE like enzyme  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 2599)  
 AUTHORS Komatsu, T., Suzuki, Y. and Sugano, S.  
 TITLE Direct Submission  
 JOURNAL Science, Virology: 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
 Submitted (25-JUL-2000) Takami Komatsu, the Institute of Medical  
 Science, Virology: 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
 (E-mail: komatsu@ims.u-tokyo.ac.jp, Tel:81-3-5449-5283(ex.75283),  
 Fax:81-3-5449-5416)  
 FEATURES  
 source  
 1. 2599  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="X"  
 /map="Xp22"  
 /clone="kala4505"  
 /clone\_lib="Sugano human cDNA library kala"  
 /dev\_stage="adult"  
 /country="Japan"  
 /note="Vector: pME18s FL3, lab host: Top10"  
 1. 2599  
 /gene="ace2"  
 55. 2472  
 /gene="ace2"  
 /codon\_start=1  
 /product="ACE2"  
 /protein\_id="BAB40370.1"  
 /db\_xref="GI:13516972"  
 /translation="MSSSSWLLSLVAATAASTTEQAKFLDKFNHAEADLFYQSS  
 LASNRYNTNTEPNVONMNNAGDKMSAFKESQTLAOMPLLEIGNLVKQLQALQD  
 NGSSVLSEDSKRLNTILNTMSTIYSGKVCNPDPNOECLLEPGNETMAISLDYNE  
 RLWAMESWRESEVKOLRPLEETVYLNKEMARAHNEYEDVDYNGDYGVDYS  
 RGQILIEVETHEFEIKFLYELHAYRAKLMAHNEYISIPICLPAHLIGDMGSEMT  
 NLSTLVPEFGCKNIDYTDAMVDQADQRIKREAKFEVSVGLPMTQGVENSMLI  
 DPGNVQKAVCHPTAMDIGKDFRLIMCTVTVMDDFTLAHENGHIQYDAVAOPFL  
 RGNANEGFHEAVEIMSLAATPKHLKSLGSLSPDQENENETFLKQALTIYGLT  
 PFTYMEKKMWMVFKGRIKPKDMKKRMKEKREIVGVPEPHDEYCPALFQVSN  
 DYSFIRYTRITLQFOFOEALCOAAKHBEGRHCDSNSTEAGOKLENNLRGKSEPW  
 TLALENVGAKNMNVRLNVEPFLTWLKDQKNSFYVSGSTDMSPYADOSTKXVLSL  
 KSALEGRAYEMNDENMYLERSVAAYARQFLVKQOMLLEGEDRYANLKRISFN  
 FETVAPAKNSDILPRTEVERKAIRMSRRIIDARLNDNSLEFLGIQPTLGPNOQPSV  
 IMLIVGVVGVIVGIVILIFTGIRDKKKRKRASGENPYASIDISKGENNPQNT  
 DDVQFSF"  
 BASE COUNT 808 a 503 c 590 g 698 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 2599  
 Score: 4288.00 Matches: 804  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.88% Mismatches: 0  
 Query Match: 99.93% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-978-385-2 (1-805) x AB046569 (1-2599)  
 QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 55 ATGTCAAGCTCTTCCCTGGCTCTTCTCAGCCTTGTCTGATACGTGCTGATCCACC 114  
 QY 21 IleglglglAlaAlaLysThrPheLeuAspLysPheAsnHISGLIAlaGluAspLeuPhe 40  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 115 ATTGAGGAACAGGCCAAGACATTTTGGACAAAGTTTACACAGCAAGCCGAGACCTGTC 174  
 QY 41 TTYGInSerSerLeuAlaSerTrpAsnTrpAsnTrpAsnTrpAsnTrpAsnValGln 60  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 175 TATCAAAAGTTCATGCTGCTTGGATTTATATACACCAATTTACTGAAAGAAATGTCCAA 234  
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnInsTrpLeuAla 80  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 235 AACATGAATTAACGCTGGGACAAATGCTCTCTTTTAAAGAACGCTCCACACTGGCC 294  
 QY 81 GlnMetTrpProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnIleGlnAlaLeu 100  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 295 CAATGTATCCACTACAGAAATTCGAATCTCACAGTCACGCTTCAAGCTGACAGGCTCT 354  
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 355 CAGCAAAATGGGCTTCAAGCTCTCAGAACAGCAAGCAAGCAAGCTTGAACCAATTCCTA 414  
 QY 121 AsnThrMetSerThrIleTrpSerThrGlyLysValLysAsnProAspAsnProGlnGlu 140  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 415 AATACATATGAGCACCATCTACAGTCTGAGAAAGTTTGTAACCCAGATTAATCCAAAGAA 474  
 QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTrpAsnGlu 160  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 475 TGCTTATTTACTTGAACAGAGTTGAATGAATATGGCAACAGTTTATGACTACAAATGAG 534  
 QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTrp 180  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 535 AGGCTCGGCTTGGGAAAGCTGAGATCTGAGTGGAGTGGCAACACTGATGAGCCATTATAT 594  
 QY 181 GlnGluTrpValValLeuLysAsnGluMetAlaArgAlaAsnHISLysGluAspTrpGly 200  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 595 GAGAGATATGCTGCTTGAATAATGAGATGGCAAGCAACATATGAGAGACTATGAG 654  
 QY 201 AspTrpTrpArgGlyAspTrpGluValAsnGlyValAspGlyLysAspTrpSerArgGly 220  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 655 GATATTTGAGAGAGAGACTATGAAATGAATGAGGTTGATGAGCTATGACTACAGCCGCCG 714  
 QY 221 GlnLeuIleGluAspValGlnHISThrPheGlnGluIleLysProLeuTrpGluHISLeu 240  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 715 CAGTTGATTTGAAGTGTGGAAACATACCTTTGAAGACATTAACCATTTATGAACATCTT 774  
 QY 241 HisAlaTrpValAlaArgAlaLysLeuMetAsnAlaTrpProSerTrpLysSerProIleGly 260  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 775 CATGCTATGAGAGGCAAAAGTTGATGAATGCTATCTCTTAATACAGTCCAAATTGGA 834  
 QY 261 CysLeuProAlaHISLeuLeuGlyAspMetTrpGlyArgPheTrpHisLeuTrpSer 280  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 835 TGCCCTCCCTGCTCATTTGCTTGTGATATGTGGGGATGATTTTGGCAAAATGCTACTCT 894  
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 895 TTGACAGTTCCTTTGGACAGAAACCAATAGATGTACTATGAGAAAGGGGAGCCAG 954  
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 955 GCCTGGATGACACAGAAATATTCAGAGAGGCGAGAAAGTCTTTGTATGTGTGGTCTT 1014  
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGluAsnValGln 340  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1015 CCTAATATGACTCAAGAGATTCTGGAAATATTCATGCAACGAGCCAGCAAAATGTTCAG 1074  
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1075 AAAGCAGTGTCCATCCACAGCTTGGGACCTGGGAAAGCCGCTTCAGAGATCTTATG 1134  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHISGluMetGlyHisIleGln 380  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1135 TGACAAAGGTGACATGAGAGCTTCTGACAGCTCATATGAGATGGGCAATATTCAG 1194



QY 381 TyrAspMetAlaTyrAlaIaGlnProPheLeuLeuArgSnglYAlaasnGluGlyPhe 400  
 |||||  
 Db 1195 TATGATATGGCATATCTGCACACACTTTTCTGCTAGAGAAAGGAGCTATATAGAGATTTC 1254  
 QY 401 HsGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
 |||||  
 Db 1255 CATGAGAGCTGTGGGGAATCATGTCACTTCTGGAGGCACACCTAAGATTATAAATCC 1314  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440  
 |||||  
 Db 1315 ATTTGGCTTCTGTCACTGCACCAATTTTCAAGAAAGCAATGAAACAGAAATTAACCTCTGCTC 1374  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysThrArg 460  
 |||||  
 Db 1375 AACAAGCACTCAGCATTTCTGGACTCTGCCATTACTTACATTGAGAGTGGAGG 1434  
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpPheGluMet 480  
 |||||  
 Db 1435 TGGATGGCTCTTAAAGGGGAATTCCTCAAGACAGTGGATGAAAAAGTGGGAGATG 1494  
 QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500  
 |||||  
 Db 1495 AAGCGAGAGATAGTGGGGTGGTGGACCTGTGCCCATGATGAAACATACGTGACCCC 1554  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520  
 |||||  
 Db 1555 GCATCTGCTTCCATGTTCTTAATATTAATCACTTCACTTCACTTCACTTCACTTCACTT 1614  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlnAlaValAlaLysHisGluGlyProLeuHis 540  
 |||||  
 Db 1615 TACCAATTCACACTTCAAGAGACACTTGTCAAGCAGCTAAACAGAGAGCCCTGTGCAC 1674  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560  
 |||||  
 Db 1675 AATATGACATCTCAAACTCAACAACTGACAGCAAGCTGCAATATCTGAGGCTT 1734  
 QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGlnAsnValAlaGlyAlaLysAsnMetAsn 580  
 |||||  
 Db 1735 GGAATAATCAGAACTGACCTGACCTTACCATTTGGAATATGTTGAGGAGCAAAACATGAT 1794  
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpPheLysAspGlnAsnLys 600  
 |||||  
 Db 1795 GTAAGGCCACTGCTCACTACTTGTGAGCCCTTATTTACCTGGCTGAAAGACCAAGAAC 1854  
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620  
 |||||  
 Db 1855 AATTTCTTTTGGATGAGTACCGACTGAGTGCATATGCAAGCAACAGCATCAAGATG 1914  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlnLysPheLysAlaTyrGlnLysAsnAspAsnGluMet 640  
 |||||  
 Db 1915 AGGATTAACCTTAATAATCAGCTCTTGGAGATAGACATATGTAATGGAACAGACATGAATG 1974  
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660  
 |||||  
 Db 1975 TACCTGTTCCATCATCTGTTGCAATGCTATGTAAGGAGCACTTTTAAAGTAAAAAAT 2034  
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 |||||  
 Db 2035 CAGATGATTTCTTTGGGAGAGAGATGTGGAGTGGCTAATTTGAAACCAAGATCTCC 2094  
 QY 681 PheAsnProPheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700  
 |||||  
 Db 2095 TTTAATTTCTTTCACATGCACTTAATAATCTGCTCATATCATCTTCTAGAACTGAAGTT 2154  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 |||||  
 Db 2155 GAAAGAGCATCAGAGATGTCCCGAGCCGATCATATATGCTTCCCTCTGAAAGAAC 2214  
 QY 721 SerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 |||||  
 Db 2215 AGCTAGAGTTTCTGGGATACAGCAAACTTGGACTCTCTAAGCAAGCCCTGTTTCC 2274

QY 741 IleTrpLeuIleValPheGlyValAlaMetClyValIleValValGlyIleValIleLeu 760  
 |||||  
 Db 2275 ATATGGCTGATTTGTTTGGAGTTGTGATGGAGTGTATGATGTCATTCATCTCGT 2334  
 QY 761 IlePheThrGlyIleArgAspArgLysLysAsnLysAlaArgSerGlyLysPro 780  
 |||||  
 Db 2335 ATCTTCACTGGGATCAGAGATCGGAGAGAGAAAAATTAAGCAGAGATGGAGAAATCT 2394  
 QY 781 TyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAsp 800  
 |||||  
 Db 2395 TATCCTTCATCATGATATAGCAAGAGAAAAATATATCAGAGATTCCAAAACATGATGAT 2454  
 QY 801 ValGlnThrSerPhe 805  
 |||||  
 Db 2455 GTTCAGACTCTCTT 2469  
 RESULT 13  
 HSM800880  
 LOCUS  
 DEFINITION Homo sapiens mRNA; cDNA DKFZp434A014 (from clone DKFZp434A014);  
 partial cds.  
 ACCESSION AL110224  
 VERSION AL110224.1 GI:5817159  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens.  
 Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 AUTHORS Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-AUG-1999) MIPs, Am Klopfersplitz 18a, D-82152  
 Martinstried, GERMANY  
 COMMENT  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing  
 consortium of the German Genome Project.  
 This clone (DKFZp434A014) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available  
 at <http://www.mips.biochem.mpg.de/proj/cDNA/>.  
 FEATURES  
 source  
 1..3341  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp434A014"  
 /issue\_type="testis"  
 /clone\_lib="434 (synonym: htes3). Vector pSport1; host  
 DH10B; sites NotI + SalI"  
 /dev\_stage="adult"  
 1..2417  
 /gene="DKFZp434A014"  
 <1..2417  
 /note="DKFZp434A014"  
 /note="strong similarity to peptidyl-dipeptidase A"  
 /product="hypothetical protein"  
 /protein\_id="CAB53682.1"  
 /db\_xref="GI:5817160"  
 /translation="SSSSWLLSLVAVTAAHSTTEQAKFTLDFKNHNEADLPFOSSL  
 ASMNNTNTEENYONMNNAGKNSAFKEOSTLAQWPLEIOELNLTAKIOLALOON  
 GSSVLEDSKSKRLNTILNMSTIYSTGVCAPNDNQBELLPEGLNETMANSLDYNER  
 LMAWMSSEVCKOIRPIYEEYVYKKNMARANHEDYGDYWRGDIYVNGVDYISR  
 GLTIDVETHELEIKLVEYHHAIVRAKLMAVPSYISPICLPAHLIGDMWGRFWN  
 IYSLIVPGQKRENDIVDAMVDADQRIKREKEKFEVSGVGLNMNGEWNESMLND  
 NGNOKAVCHPFAMDLGGDFRIILKCTVTVDLELAHEMGHIOYMAVYAOPELLR  
 PANGVGFHEAVGEIISLAAPPKHLKSGLSPPDEENETELNEFLKQALTTVGTLP  
 FTYMEKRMVVFGEIPKDDMMKKMKMKREIIVGYVPPVPHDEYCDPASLFFVSD  
 YSTIYVYRTILYQRFQALCOAKHBEPLKDCISNTEAGQLEMLRLGKSEVNT  
 LALENVYAKMNNVRFPLNIFEPLETLWDKNKSPGWSWDSPYADQSLKVISLK  
 SALGDKAYEAMNDENMYLFRSSVAYAMROYFLKVNOMTLFGEEDVYRANLKPRISENF  
 FVTAKNVSDIIPREVEKAIKMRSRINDAFRLNDNSLEFLGIQPTLGPPNQPVSIF

WLIVGVGMIVIVGIVILLIFTGIRDRKKKKKARSNGENPYASIDISGENNPGFQND  
 DVQSF\*  
 polyA\_signal 3266 .3271  
 polyA\_site 3281  
 BASE COUNT 1058 a 630 c 740 g 913 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 0 Length: 3341  
 Score: 4281.00 Matches: 803  
 Percent Similarity: 99.88% Conservative: 0  
 Best Local Similarity: 99.88% Mismatches: 1  
 Query Match: 99.77% Indels: 0  
 DB: Gaps: 0

US-09-978-385-2 (1-805) x HSNM00880 (1-3341)

QY 2 SerSerSerSerTrpleuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThrIle 21  
 |||||  
 Db 3 TCAGGCTCTTCTGCTCTCTCTCAGCCCTGTGGCTGTAAGCTGCTCAGCCACCAT 62  
 QY 22 GlnGlnGlnAlaAlaSerThrPheLeuAspLysPheAsnHisGlnAlaGlnAspLeuPheTyr 41  
 |||||  
 Db 63 GAGGACAGAGCCAGACATTTTGGCAAGTTTAACCCAGAGCCGAGACCTGTTCTAT 122  
 QY 42 GlnSerSerLeuAlaSerTrpAsnTrpAsnThrAsnIleThrGlnGlnAsnValGlnAsn 61  
 |||||  
 Db 123 CAAAGTCACTGCTCTTGGAATTTATACACCAATATTACTGAAGACATGTCACAAAC 182  
 QY 62 MetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAlaGln 81  
 |||||  
 Db 183 ATGATTAATAGCTGGGAGCAATGCTGCTTTTAAAGAACAGCTCCACACTTGCCCA 242  
 QY 82 MetTrpProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeuGln 101  
 |||||  
 Db 243 ATGTATCCACATCAAAATAATCAGATTCACAGTCACAGTCACCTGAGGCTTTCAG 302  
 QY 102 GlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrIleLeuAsn 121  
 |||||  
 Db 303 CAAATAGGCTCTTCACTGCTCTCAGAAAGCAAGACAAACGGTTGACACATTTCTAAT 362  
 QY 122 ThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlyCys 141  
 |||||  
 Db 363 ACAATGAGCACCATTTACAGTACTGGAAAAGTTTGTAACCCAGATTAATCCAAAGATGTC 422  
 QY 142 LeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTrpAsnGlnLys 161  
 |||||  
 Db 423 TTATTACTTGAAACGAGTTTGAATGAATATGACAAACAGTTTACATACATGAGAG 482  
 QY 162 LeuTrpAlaTrpGlnSerTrpArgSerGlnValGlyLysGlnLeuArgProLeuTrpGln 181  
 |||||  
 Db 483 CTGTGGCTGGGAAAGCTGAGATCTGAGTGGCAACAGCTGAGGCGCATTAATATGA 542  
 QY 182 GlnTrpValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspTrpGlyAsp 201  
 |||||  
 Db 543 GAGTATGTGGTCTTGAATAATGAGATGCGAAGCAAAATCATTTATGAGACATATGGGGAT 602  
 QY 202 TyrTrpArgGlyAspTrpGlnValAsnGlnValAspGlyTyrAspTrpSerArgGlyGln 221  
 |||||  
 Db 603 TATTTGAGAGAGACTATGAAATGGGATGATGCTATGACTACAGCGCGGCGAG 662  
 QY 222 LeuIleGlnAspValGlnHisThrPheGlnGlnIleLysProLeuTrpGlnHisLeuHis 241  
 |||||  
 Db 663 TTGATTCAGACATGTCGACATACCTTTGAAGAGATTAACCATTAATATGACATCTTCAT 722  
 QY 242 AlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrLysSerProIleGlyCys 261  
 |||||  
 Db 723 GCTATGTGAGGCAAAAGTATGATGAATGCTATCTCTATATCGTGCATTTGGATGC 782  
 QY 262 LeuProAlaHisLeuLeuGlnGlyAspMetTrpGlyArgPheTrpHisLeuTrpSerLeu 281  
 |||||  
 Db 783 CTCCTGCTCATTTGCTGTGATATGTGGGTAGATTTTGGCAAAATCTGTACTCTTTG 842

QY 282 ThrValProPheGlnGlnLysProAsnIleAspValThrAspAlaMetValAspGlnAla 301  
 |||||  
 Db 843 ACAGTTCCTTTGGACGAGAAACCAACATATGATTAATGATCATATGATGATGGAGCAGGCG 902  
 QY 302 TrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerValGlyLeuPro 321  
 |||||  
 Db 903 TGGGATCCACAGAAATATTCAAAGAGGCGGAGAAAGTTCTTTGATCTGTGGCTTCTCT 962  
 QY 322 AsnMetThrGlnGlnPheTrpGlnAsnSerMetLeuThrAspProGlnAsnValGlnLys 341  
 |||||  
 Db 963 AATATGACTCAGAGATTTCTGGAAATATTCATGCTATACGAGACCCAGAAATGTTCCGAAA 1022  
 QY 342 AlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMetCys 361  
 |||||  
 Db 1023 GCAGTCGCCATCCACAGCTGGAGCTGGGAGAGGCGACCTTCAGAGATCTTATGTGC 1082  
 QY 362 ThrLysValThrMetAspAspPheLeuThrAlaHisIleGlnMetGlnHisIleGlnTyr 381  
 |||||  
 Db 1083 ACAAGGTGACAAATGAGACGACTTCTGACAGCTCATCATGATGAGTGGGCAATCCAGTAT 1142  
 QY 382 AspMetAlaTyrAlaIleAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPheHis 401  
 |||||  
 Db 1143 GATATGGCATATGCTGCACAAACCTTTCTGCTAAGAAATGAGACTAATGAAAGATTCAT 1202  
 QY 402 GlnAlaValGlyGlnIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSerIle 421  
 |||||  
 Db 1203 GAAGCTGTGGGAAATCATGTCACTTCTGACCCACCTTAAGCATTAATAATCCAT 1262  
 QY 422 GlyLeuLeuSerProAspPheGlnGlnLysAspAsnGlnIleThrGlnLysAsnPheLeuLys 441  
 |||||  
 Db 1263 GGTCTTGTGCACCCGATTTTCAAGAGACAAATGAACAAATAAATTTCTGCTCAAA 1322  
 QY 442 GlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTrpArgTyr 461  
 |||||  
 Db 1323 CAAGCAGCTCAGCATTTGTTGGAGCTCTGCCATTTACTTACATGTTAGAGAGTGGAGGTG 1382  
 QY 462 MetValPheLysGlyGlnIleProLysAspGlnTrpMetLysLysTrpTrpGlnMetLys 481  
 |||||  
 Db 1383 ATGCTCTTAAAGGGAATCCCAAGACAGAGGATGAAAAGCTGTGGGAGATGAG 1442  
 QY 482 ArgGlnIleValGlyValGlnProValProHisAspGlnIleThrTyrCysAspProAla 501  
 |||||  
 Db 1443 CGAGAGATATGGGAGGTGGTGAACCTGCGCCATGATGAACATCTGTGACCCGCA 1502  
 QY 502 SerLeuPheHisValSerAsnAspLysSerPheIleArgTyrTrpThrArgThrLeuTyr 521  
 |||||  
 Db 1503 TCTGCTTCATGTTTCAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 1562  
 QY 522 GlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHisLys 541  
 |||||  
 Db 1563 CAATTCAGTTTCAAGAGCACTTGTGCAAGCACCTAAACATGAAAGCCCTCTGCAACAA 1622  
 QY 542 CysAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGly 561  
 |||||  
 Db 1623 TGTGACATCTCAAACTCTACAGAAAGCTGAGAGAAAGCTTTCAATATGCTGAGGCTTGA 1682  
 QY 562 LysSerGlnProTrpThrLeuAlaLeuGlnAsnValGlyAlaLysAsnMetAsnVal 581  
 |||||  
 Db 1683 AAATAGAAACCTGGACCTGACCTGATGAAATATGTTAGAGAGCAAGAACATGAATGTA 1742  
 QY 582 ArgProLeuLeuAsnTrpPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLysAsn 601  
 |||||  
 Db 1743 AGGCACAGCTCAACTACTTGTGACCCCTTATTTACTCTGCGGAGAAAGCCAGAAACAAAT 1802  
 QY 602 SerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysValArg 621  
 |||||  
 Db 1803 TCTTTTGGGATGGAGTACGACTGAGTCCATATGCAAGCCAAAGCATCAAAAGTACG 1862  
 QY 622 IleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGlnTrpAsnAspAsnGlnMetTyr 641  
 |||||  
 Db 1863 ATTAACCTTAATAATACGCTTGTGAGATTAACCATATGAAATGAGACGAAATGAATGTAC 1922  
 QY 642 LeuPheArgSerSerValAlaTyrAlaMetArgGlnIleThrPheLeuLysValLysAsnGln 661

```

|||||
Db 1923 CTGTTCCGATCATCTGTTCCATATGCTATGAGGCGAGTACTTTTAAAGTAAATAATCAG 1982
QY 662 MetileLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgLleSerPhe 681
Db 1983 ATGATCTCTTTTGGGAGAGAGATGGGAGTGGCTAAATTTGAAACCAAGAAATCTCTTT 2042
QY 682 AsnPheValThrAlaProLysAsnValSerAspLleLeuProArgThrGluValGlu 701
Db 2043 AATTTCTTGTCTGACGACCTAAATAATGCTGATATCATCTCCAGAACTAAATTGAA 2102
QY 702 LysAlaIleArgMetSerArgSerArgLleAsnAspAlaPheArgLeuAsnAspAsnSer 721
Db 2103 AAGGCATAGAGATGTCCTGGAGCCGCTATCATGATGCTTCCGCTGTAATGACAACAGC 2162
QY 722 LeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerLle 741
Db 2163 CTAGAGTTCTTGGGATACAGCCACACTTGGACCTCTTACCAACGCCCCGTTTCCATA 2222
QY 742 TrpLeuIleValPheGlyValValMetGlyValIleValIleValIleValIleLeu 761
Db 2223 TGGCTGATTTGTTTGGAGTTGTGATGGAGATAGTGGTTGGCATTTGTCATCTGATC 2282
QY 762 PheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnProTyr 781
Db 2283 TTCACATGGGATCAGAGATCGAGAGAGAAATAAAGCAAGAGTGGAGAAATCTTAT 2342
QY 782 AlaSerIleAspLleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAspVal 801
Db 2343 GCCTCATCGATATTAGCAAGAGAAATATCATCAGATTCACAAACACTGATGATGTT 2402
QY 802 GlnThrSerPhe 805
Db 2403 CAGACCTCTCTT 2414

RESULT 14
AX463938
LOCUS AX463938 3732 bp DNA linear PAT 16-JUN-2002
DEFINITION Sequence 71 from Patent WO0140466.
ACCESSION AX463938
VERSION AX463938.1 GI:21898986
KEYWORDS
SOURCE human.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,
Gao,W.Q., Gerlitsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,
Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,
Wood,W.L. and Zhang,Z.
Secreted and transmembrane polypeptides and nucleic acids encoding
TITLE
same
JOURNAL Patent: WO 0140466-A 71 07-JUN-2001;
Genentech Inc. (US)
FEATURES
source location/Qualifiers
1..3732
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1137 a 722 c 821 g 1052 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3732
Score: 4142.00 Matches: 802
Percent Similarity: 85.33% Conservatave: 2
Best Local Similarity: 85.23% Mismatches: 1
Query Match: 96.53% Indels: 138
DB: 6 Gaps: 1
US-09-978-385-2 (1-805) x AX463938 (1-3732)
QY 1 MetSerSerSerThrLeuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20

```

```

|||||
Db 40 ATGTCACACTTTCTTGGCTCTCTCTCAAGCCTTCTTCTGTAATGCTGCTCACTCCACC 99
QY 21 IleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
Db 100 ATTGAGAGACAGGGCCAGACATTTTGGACAGTTTAAACCAAGCCAGAGACCTGTC 139
QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGln 60
Db 160 TATCAAAAGTTCACTTGGCTTCTTGGAAATTAACCCAAATTAATCACTGAAGAGAAATGCCAA 219
QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 220 AACATGAAATATGCTGGGGCAAAATGCTGCTTTTAAAGAGACGTCACACTTGGC 279
QY 81 GlnMetYrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
Db 280 CAAATGTATCCACACACAAAGAAATTCAGAAATCTCAGACTCAAGCTTCAGCTGACGCTCTT 339
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
Db 340 CAGCAAAATGGGCTTCAAGTCTCTCAGAAAGACAGCAAGCGTTGAACACAAATTCCTA 399
QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 400 AATTCATATGAGCACACCATCTACAGTACGAAAGTTTGTAACCCACATATATCCACAAGAA 459
QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 460 TGCTTATTAATCTTAAACAGGCTTGAATGAATATGCAACAGTTTACATCTCAATGAG 519
QY 161 ArgLeuThrPalatTrpGlnSerTrpArgSerGluValGlyGlnLeuArgProLeuTyr 180
Db 520 AGGCTCTGGGCTTGGGAAACCTGGAATCTGAGGTGGCAAGCAGCTGAGCCATTATAT 579
QY 181 GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
Db 580 GAAGAGATATGCTCTTGAATAATGAGATGGCAAGCAACAAATCATATATAGAGCATATGGG 639
QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
Db 640 GATTATTTGAGAGAGACATATGAAATGAATGGGTGATGCTATGATCAGTACAGCGCGC 699
QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240
Db 700 CACTGATTTGAAATATGAGACATATCTTGAAGAGATTTAAACATTATATATACACTCTT 759
QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 760 CATGCTATGTGAGGCAAAAGTTGATGAATGCTATCTCTATATATCAGTCCAAATTTGA 819
QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheThrAsnLeuTyrSer 280
Db 820 TGGCTCCCTGCTCATTTGCTTGGTATATGTGGAGATTTTGGACAAATCTGTATCTCT 879
QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 880 TTGACACAGTTCCCTTTGGACAGAAACCAACATAGATTTACATGATGCAATGTTGACACAG 939
QY 301 AlaTrpAspAlaGlnArgLlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
Db 940 GCGTGGGATGACAGAGAAATATCAAGAGCGCGGAAATCTTGTATCTGTGCTGCTCT 999
QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyLysAsnValGln 340
Db 1000 CCTAATATATACCAAGATATTCGGGAAATTTCACTATGACGAGCCAGGAAATGTTAG 1059
QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1060 AAAGAGATCTGGCATCCACAGCTTGGGAGCTGGGAGAGGGGAGACTTCAGATCTTATG 1119
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380

```



```
PC      A61K37/02
CC
FH      Key
FT      Location/Qualifiers
FT      source
          1. .2262
FEATURES
source   Location/Qualifiers
          1. .2262
```

BASE COUNT	693 a	450 c	523 g	596 t
ORIGIN				

Alignment Scores:

Pred. No.:	2,28e-305	Length:	226
Score:	3740.50	Matches:	711
Percent Similarity:	90.57%	Conservative:	9
Best Local Similarity:	89.43%	Mismatches:	12
Query Match:	87.17%	Indels:	63
DB:	6	Gaps:	3

US-09-978-385-2 (1-805) x E39034 (1-2262)

QY	11	LeuValAlaValThrAlaIaGlnSerThrIleGluGluGlnAlaIysThrPheLeuAsp	30
Db	64	CTTGTGCTGTAAGTCTGCTGCTCACTCCACATTGAGGACAGCGCAAGACATTTTGGAC	123
QY	31	LysPheAsnHisGluAlaGluAspLeuPheGlyGlnSerLeuAlaSerTPAsnTyr	50
Db	124	AAAGTTTAACCCAGCAAGCCGAAAGACTGTGTCATCAAGTTCACTGCTTCTTGGAAATTAT	183
QY	51	AsnThrAsnIleThrGluGluAsnValGlnAsnMetAsnAsnAlaGluAspLysTrpSer	70
Db	184	AACCCACATATTACTGGAAGAGAAATGTCAAAACATGATATATGCTGGGGCAAAATGCTCT	243
QY	71	AlaPheLeuLysGluGlnSerThrLeuAlaGlnMetThrProLeuGlnGluIleGlnAsn	90
Db	244	GCCCTTTTAAAGGACACAGTCACACTTGCCCAAAATGTATCCACTACACAGAAATTCACAAT	303
QY	91	LeuThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerValLeuSerGlu	110
Db	304	CTCACAGTCAAGCTTCAGCTGCAGAGGCTCTTCAGCAAAATGGGCTTCAGTGCCTCGAA	363
QY	111	AspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGly	130
Db	364	GACAGAGCAACGGTTGGAACACAAATCTAAATACATGACGACACATCTACAGTACTGGA	423
QY	131	LysValCysAsnProAspAsnProGlnGlnCysLeuLeuLeuGlnLysProGlyLeuAsnGlu	150
Db	424	AAAGTTGTATACCCACAGTATATCCACAGAAATGCTTATCTTACCCACAGGTTTGAATGA	483
QY	151	IleMetAlaAsnSerLeuAspTyrAsnGluArgLeuTrpAlaTrpGlnSerTPArgSer	170
Db	484	ATATATGCCAAACAGTTTAGACTACATAGACAGGCTGTGGGCTTGGGAAACCTGGAGATCT	543
QY	171	GluValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGluMet	190
Db	544	GAGGTGGCAGAGAGCTGAGGCCATATATGAAGAGATGTGCTTGAATAATGAGATG	603
QY	191	AlaArgAlaAsnHisTyrGluAspTyrGluLysTyrTrpArgLysAspTyrGluValAsn	210
Db	604	GCAAGACCAATATATATGAGACTATAGGGGATATATGAGAGCAACATATGAAGATTAAT	663
QY	211	GlyValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluLysPylGluHisThrPhe	230
Db	664	GGGGTAGATGGCTATGACTACAGCGCGGCCAATTGATTAAGATGTGGACATTAACCTTT	723
QY	231	GluGluIleLysProLeuTyrGluHisSleuHisAlaTyrValArgAlaLysLeuMetAsn	250
Db	724	GAGAGACATTAAACATTATATGAACATCTTCACCTATGTGAGGGCAAAAGTTGATGAAAT	783
QY	251	AlaTyrProSerTyrIleSerProIleGlyCysLeuProAlaHisSleuLeuGlyAspMet	270
Db	784	GCTATTCCTTCATATCATGTCCAAATGGAGGCTCCGCTGCATATTTGGTTGGTATATG	843

QY	271	TRPGLYATGPEHETPRThAspLeuTyrSerLeuThrValProPheGlyGlnLysProAsn	290
Db	844	TGGGGTACGATTTTGGACAAATCTGACTCTTGGACAGTTCCCTTTGGACAGAAACCAAAAC	903
QY	291	ILesapValThrAspAlaMetValAspGlnAlaTrpAspAlaGlnArgILEPheLysGlu	310
Db	904	ATAGATGTGTACTGATGCAANTGGTGGACAGCCCTGGGATGCACAGAAATATTCAGAGAG	963
QY	311	AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnArgPheTrpGluAsn	330
Db	964	GCCGGAAGTCTTTGTATCTGTTGGCTTCTTAAATGACTGAAGGATTCGGGAAAT	1022
QY	331	SerMetLeuThrAspProGlyAsnValGlnLysAlaValCysHisIserProThrAlaTrpAsp	350
Db	1024	TTCATGCTCAAGGAGACCCAGCAAAATGGTTCAGAAACACATCTGCCATCCACAGCTTGGAC	1083
QY	351	LeuGlyLysGlyAspPheArgILEuMetCysThrItyValThrMetLysPheLeu	370
Db	1084	CTGGGGAAGGGGAGCTTCAGGATCTTATGGCAAAAGGTGCAATGAGACACTTCCG	1144
QY	371	ThrAlaHisHISgluMetGlyHisILEGlnTyrAspMetLarValAlaGlnProPhe	390
Db	1144	ACAGCTCATCATGAGATGGGCGCATATCCAGATGATATGGCATATGCTGCACACCTTTT	1203
QY	391	LeuLeuArgAsnGlyAlaAsnGlyLysPheHisGluAlaValGlyIleMetSerLeu	410
Db	1204	CTGGTAAAGAAATGGAGCTCAATGAAGATTCATCAAGACCTGTGGGAAATCATATGCACAT	1263
QY	411	SerAlaAlaTrpProLysHisLeuLysSerILEGlyLeuLeuSerProAspPheGlnGlu	430
Db	1264	TCTGAGCCACACCTAAGACATTTAAATTCATTTGGTCTTCTGTGACACCGATTTTCAGAA	1322
QY	431	AspAsnGluThrGluILEasnPheLeuLeuLysGlnAlaLeuThrILEValGlyThrLeu	450
Db	1324	GACCATGAAACAGAAATTAATCTCTGCTCAACAGACCTCCGATTTGTGGACTCTG	1383
QY	451	ProPheThrTyrMetLeuGluLysTrpArgTrpMetValPheLysGlyIleProLys	470
Db	1384	CCATTACTTTCATGTTAGAGAAATGGAGGTGGATGTTCTTTAAAGGGAATTTCCCAA	1443
QY	471	AspGlnTrpMetLysLysTrpTrpGluMetLysArgGluILEValGlyValGluPro	490
Db	1444	GACCATGTGAGAAAGATGGTGGAGATGAAA-----	1476
QY	491	ValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAspTyr	510
Db	1476	-----	1476
QY	511	SerPheILEArgTyrTrpThrArgThrLeuTyrGlnPheGlnAlaLeuLysCys	530
Db	1477	-----TATTCACACAGACCCCTTTACCATTCACAGTTTCACAGACACTTTGT	1524
QY	531	GlnAlaAlaLysHisGluGlyProLeuHisLysCysAspLIEserAsnSerThrGluAla	550
Db	1535	CAGAGAGCTTAAACCTGGAAGCCCTCTGCACAAATGTGCACTCTCAAACTTCACAGAACT	1584
QY	551	GlyGlnLysLeuPheAsnMetLeuArgGlyGlyLysSerGlnProTrpThrLeuAlaLeu	570
Db	1585	GGACAGAAACGTTCATATGCTGAGGCTTGGAAATTCAGAACCCCGACACCTAGACTTG	1644
QY	571	GluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGluPro	590
Db	1645	GAATATGTTGTAGAGAGCAAGAACATGATGTAAAGCCACTGCTCAACTACTTGAGCC	1704
QY	591	LeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGlyTrpSerThrAspTrp	610
Db	1705	TTATTATTCCTGGCTGAAAGACCAGAACAGAAATCTTTTGTGGATGGAGTACGACTGG	1764
QY	611	SerProTyrAlaAspGlnSerILELysValArgILESerLeuLysSerAlaLeuGlyAsp	630
Db	1765	AGTCCATGG-----	1773





Db	946	TTGACTGTTCCCTTTCACAGAAACCAACATGATGTTACTGATGACATATGATGATCAG	1005
Qy	301	AlATrPAsPAlAGlInArGllIePheLysGluAlaGluLysPhePheValSerValGlyLeu	320
Db	1006	GGCTGGGAGCGAAGAAAGATATTTCAGAGGAGAGAAATCTTTGTTCTGTGGCCTT	1065
Qy	321	ProAsMeTrhGlnGlyHeTrpGluAsnSerMeTrLeuTrhAspProGlyAsnValGln	340
Db	1066	CCTCATATGATCAAGAGATTCTGGGCAAACTATGCTGACTGAGCAGCAGATGGCCGG	1125
Qy	341	LysAlaValAlcYHisProThralATrPAsPLeuGlyLysGlyAsPheArgIleLeuMet	360
Db	1126	AAAGTTTGTGGACCCACACCTTGGGATCTGGGACAGGAGACTTCAGAAATCAAGATG	1185
Qy	361	CysTrhLysValTrhMeTrAsPAsPheLeuTrhAlaHisHISGluMetGlyHisIleGln	380
Db	1186	TGTACAAAGCTCAATGACATGGACACTCTTGTGACAGCCCTCCGAGATGGGACATCCAA	1245
Qy	381	TyrAsPMeTrAlATrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyLysPhe	400
Db	1246	TATGACATGGCATATGCCAGGCAACCTTCTCGTCAAGAAAGGAGCCATGAAGGTTTC	1305
Qy	401	HISGluAlaValAlcYLeuIleMeTrSerLeuSerAlaAlaTrhProLysHisLysLysSer	420
Db	1306	CATGAAGCTGTGGAGAAATCATGCTCTTCTGACGTACCCCAAGCACTGCAAAATTC	1365
Qy	421	IleGlyLeuLeuSerProAsPheGlnGluAsnGluTrhArgIleAsnPheLeu	440
Db	1366	ATTGGTCTTTCGCCATCCGATTTTCAAGAAATAGCCAAACAGAGATAAACTTCCTACTG	1425
Qy	441	LysGlnAlaLeuTrhIleValAlcYTrhLeuProPheTrhTyrMeLeuGluLysTrpArg	460
Db	1426	AAACAGGATTCACATATGTGGAAACACTACCGTTTACTTACATGTTAGAGAGTGGAGG	1485
Qy	461	TrpMeTrValPheLysGlyGluIleProLysAsPglTrhPheLysTrpPheLysMet	480
Db	1486	TGGATGGTCTTTCGGGGGTGAATTTCCAAAGAGAGAGTGTGAATAAGTGGTGGGAGATG	1545
Qy	481	LysArgGluIleValAlcYValAlcYLeuProValProHisAsPglTrhTrhTyrCysAsPro	500
Db	1546	AAGCGGGAGATCGTGGTGTGGAGCCTCTGCCATGATGAACATATCTGTGACCTT	1605
Qy	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgIlyTrhArgTrhLeu	520
Db	1606	GCATCTGCTTCCATGTTTCTTAAAGATTAACCATCTTGTGAATTAACACAAAGACCATT	1665
Qy	521	TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaHisGlnGluProLeuHis	540
Db	1666	TACCAATTCAGTTTCAAGAAAGCTTGTTCAGAGCAGCTAAGATATATAGTTTCTTGCCAC	1725
Qy	541	LysCysAsPglIleSerAsnSerTrhGluAlaGlyGlnLysLeuPheAsnMeLeuArgLeu	560
Db	1726	AAATGTGACATCTCAAAATTCACATGGAAGCTGGGACAGAAAGTTCCTCAAGATGCTGACTT	1785
Qy	561	GlyLysSerGluProTrpTrhLeuAlaLeuGluAsnValAlcYValAlaLysAsnMetAsn	580
Db	1786	GGAATTCAGAGCCCTTGACCAAAAGCCTTGAAATAATGTGTAGACACAGATATGAGAT	1845
Qy	581	ValArgProLeuLeuAsnTrhPheGluProLeuPheTrhTrhPheLysAsPglAsnLys	600
Db	1846	GTAACACACCTGCTCAATTACTTCCAAACCTGTTGACCTGAAAGAGACAGCAACAGA	1905
Qy	601	AsnSerPheValGlyTrpSerTrhAspTrpSerProTyrAlaAspGlnSerIleLysVal	620
Db	1906	AATTCCTTTTGTGGGTGAGACACTGATGAGGCCCTATGCGCAGCAAAAGCTTAAAGTG	1965
Qy	621	ArgIleSerLeuLysSerAlaLeuGlyAsPlyAlaTyrGluTrhAsnAsnGlnMet	640
Db	1966	AGGATTAAGCTTAATAATGAGCTTGTGAGAGCTAATGATGATGAGACCAACACCAATG	2025
Qy	641	TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTrhPheLeuLysValLysAsn	660
Db	2026	TTTCTGTTCCGATTCATCTGTTGCATATGCCATGAGAAAGTATTTTTCATAATCAACAAAC	2085
Qy	661	GlnMeTrIleLeuPheGlyGluGluAsnValArgValAlaAsnLeuLysProArgIleSer	680
Db	2086	CAGACAGTTCCTTTCTTACAGAGATATACGATGACGACATTTGAACCAAGAGTCTCC	2145
Qy	681	PheAsnPhePheValTrhAlaProLysAsnValSerAspIleIleProArgTrhGluVal	700
Db	2146	TTTCTACTTCTTCTCCACTCAACCCAAATGTGTCTGATGTCATCTCTAAGAGTGAAGT	2205
Qy	701	GluLysAlaIleArgMeSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn	720
Db	2206	GAAAGATGCCATCAGATGTTCTTCGGGCCGCAATCATATGTCCTTGGCTGATGATGAAC	2265
Qy	721	SerLeuGluPheLeuGlyIleGlnProTrhLeuGlyProProAsnGlnProProValSer	740
Db	2266	AGCCTGGAGTTTCTGGGATTCACCCAAACACTTGAGCACCTTAACAGCCTCTGTCCAC	2325
Qy	741	IleTrpLeuIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeu	760
Db	2326	ATATGCTGATTAATTTTGGTGTGTGATGACACGTGTAAGTGTGTCATCATCATCTCG	2385
Qy	761	IlePheTrhGlyIleArgAsPArgLysLysLysAsnLysLysLysArgLysSerGlyLysAsnPro	780
Db	2386	ATTGTCACTGGGATCAAAAGTCGAAGAAGAAAAATATAACAAAGACAGAGAAACCTT	2445
Qy	781	TyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnTrhAspAsp	800
Db	2446	TATGACTGATGAGCATGTGAAGAGAAACAAATGACAGATTCCAAACACAGTATGAT	2505
Qy	801	ValGlnTrhSerPhe	805
Db	2506	GCTCAGACTTCTTT	2520
RESULT 17			
LOCUS	BC026801	2739 bp	mRNA linear ROD 07-NOV-2002
DEFINITION	Mus musculus	RIKEN cDNA 2010305105 gene, clone MGC:25940	
ACCESSION	IMAGE:4236529	mRNA, complete cds.	
VERSION	BC026801		
KEYWORDS	BC026801.1	GI:20071163	
SOURCE	MGC.		
ORGANISM	house mouse.		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
JOURNAL	1 (bases 1 to 2739)		
REMARK	Strausberg, R.		
COMMENT	Direct Submission		
	Submitted (04-APR-2002) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		



FEATURES analysis, similarity but not identity to protein.  
 source Location/Qualifiers  
 1. 2739

/organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /map="FVB/N"  
 /clone="MGC:25940 IMAGE:4236529"  
 /tissue.type="Kidney, normal, 5 month old male mouse."  
 /clone\_lib="NCI CGAP\_Kid14"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 176.2593

## CDS

/codon\_start=1  
 /product="Riken cDNA 2010305L05 gene"  
 /protein\_id="AAH26801.1"  
 /db\_xref="gi:20071164"  
 /db\_xref="locusid:70008"  
 /translation="MSSSSWLLSLVAVTTAQSITLSENAKTFLENNQEAEDISTOSS  
 LASNYNTNITEENAKKSEAAAKSAFYEOSTQAFSLQETQPIKROQLALQO  
 SSSALSDKKKOLNTLNTMTSTYTGKCNPNPOECLEBGLDEIMASTDYNS  
 RLMAEGRAEVGQOLRPLYEYVVLKEMARANNNDYGDYEDAEAGADYNN  
 RNOLIEDETFEAEIKPLYEHLHAYVRKLMIDTYSYISPGCLPAHLGGMASMT  
 NLYPLTFACKPNTDYDAMNGMDAERFQEAERFVSVGLPHMTOGPMASMTL  
 EPADGRKVCAPRTAMDGDFRKTCKTYMDNPLTAHHEMGHIQYDMAARQPTL  
 RENGAEVGEHVAETMSLSAATPKHLSTGLPDSFQDSETEINFLKQALITVGL  
 PPTVLEKRWMMVRGELPKQMKKWMKRELVGVEPLHDEYICDPSLEHVN  
 DYSFIRYRTIYQFOQALCOAKYKSLHKDISSTAGOKLKM,SLGSEBW  
 TKALENNVAGARMQVPLPLNFOPLEDFLQONNSQNTGEMSPADOSIKYRIL  
 KSALENNVAGARMQVPLPLNFOPLEDFLQONNSQNTGEMSPADOSIKYRIL  
 EFVTSPOVSVIRSEVEDAIRMSRGINDVPLNLSLEIGLHPLEPPQPPVT  
 IWLIEGVVMAVVVGIITLVTGKGRKKKRENPIDSDIGKESNAQFONS  
 DDAQTSF"

BASE COUNT 832 a 577 c 620 g 710 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,13e-291 Length: 2739  
 Score: 3579.00 Matches: 661  
 Percent Similarity: 89.57% Conservative: 60  
 Best Local Similarity: 82.11% Mismatches: 84  
 Query Match: 83.41% Indels: 0  
 Gaps: 0

US-09-978-385-2 (1-805) x BC026801 (1-2739)

QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20  
 Db 176 ATGTCAGCTCCTCGCTGCTTCCACGCTTGTTGCTGCTACTAGCTCCTCAGTCCCTC 235  
 QY 21 TleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGlnLysPhe 40  
 Db 236 ACCGAGAGAAATGCCAAGACATTTTAAACACTTAAACAGGAGCTTAAAGACCTGCTC 295  
 QY 41 TyrGlnSerSerLeuLeuSerTrpAsnLysTrpAsnLysThrAsnLysThrGluGlnValGln 60  
 Db 296 TATCAAAAGTTCCTCTCTGCTTGAATTTAATTAATACTAATACTAAGAAAATGCCCA 355  
 QY 61 AsnMetAsnAsnAlaGlnLysPheLysTrpSerAlaPheLeuLysGlnLysSerThrLeuAla 80  
 Db 356 AAGATGATGAGGCTGCACCAAAATGCTGCTCTTTATGAGAAGACAGCTAAGACTGCC 415  
 QY 81 GlnMetLysTrpLeuGlnLysLysGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100  
 Db 416 CAAAGTTTCTCCTACCAAAATCCAGACTCCGATCATCAAGCTCACTAAGAGGCCCTT 475  
 QY 101 GlnGlnAlaGlnLysSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120  
 Db 476 CACCAAAAGTGGCTCTCAGCACTCAGACAGAAACAGAGTGGTGAACAATTCG 535  
 QY 121 AsnThrMetSerThrIleLysSerThrGlyLysValLysAsnProAspAsnProGlnGln 140  
 Db 536 AACACCATGACACATTAAGTACTGAGAAAAGTTGCAACCCAAAGAACCCACAGAGA 595

QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGluLysMetAlaAsnSerLeuAspTrpAsnGln 160  
 Db 596 TGCTTATTACTTATGAGCCAGATTTGATGAATATATGGCCAGACAGACTCAACCTC 655  
 QY 161 ArgLeuTrpAlaTrpGlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuLys 180  
 Db 656 AGGCTCGGGCATTGGGAGGGCTGGAGGGCTGAGAGTGGCAAGAGCTGGAGCCGTGTAT 715  
 QY 181 GlnGluTrpValValLeuLysAsnGlnMetAlaArgAlaAsnHisTrpGluAspTrpGly 200  
 Db 716 GAGAGATATGCTCTCCGAAAACGAGATGGCAAGACCAACATTAACGATCTGGG 775  
 QY 201 AspTrpTrpArgLysAspTrpGluValAsnGlyValAspGlyLysTrpSerArgLys 220  
 Db 776 GATTATTGAGAGAGGAGCATATGAAGCAGAGGAGCAGATCAACATACTAACCCTAAC 835  
 QY 221 GlnLeuLysGlnAspValGlnHisThrPheGlnGlnLysProLeuLysTrpGlnHisLeu 240  
 Db 836 CAGTTGATGAAAGTGAAGACGATCTCGCAGAGATCAAGCATGATGAGCATCTT 895  
 QY 241 HisAlaTrpValArgAlaLysLeuMetAsnAlaLysProSerTrpIleSerProIleGly 260  
 Db 896 CATGCTATGAGAGAGAGGAGTGTGATGATGATACCTCCCTCACTACACCCCACTGGA 955  
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyLysArgPheTrpHisLeuLysSer 280  
 Db 956 TGCTCTCCCTGCCATTTGCTGGTGTGATATGAGGATGATTTGGCAAAATCTGTACCT 1015  
 QY 281 LeuThrValProPheGlyGlnLysProAsnLysAspValThrAspAlaMetValAspGln 300  
 Db 1016 TTGACTGTCTCCTTTGGACAAACCAACATATGATTTACTGATGAATGATGATCAG 1075  
 QY 301 AlaTrpAspAlaGlnArgLysPheLysGluAlaGlnLysPhePheValSerValGlyLeu 320  
 Db 1076 GCGTGGATGCAAGAAAGATATTCAGAGAGCAGAGAAATCTTCTGTTGCTT 1135  
 QY 321 ProAsnMetTrpGlnLysPheTrpGluAsnSerMetLeuThrAspProGlnValGln 340  
 Db 1136 CCTCATATGATCAGAGATTTCTGGCAACCTTATGCTGATGAGCCAGCAGATGGCCGG 1195  
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgLysLeuMet 360  
 Db 1196 AAGTGTGCTGCCACCCACAGCTTGGGATCTGGGAGACGAGACCTCAGAAATCAGATG 1255  
 QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisHisGlnMetGlyHisLysGln 380  
 Db 1256 TGTCAAAGGTCACAAATGACAACTCTTGACACCCATCAGCAGATGGAGACATCCAA 1315  
 QY 381 TyrAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400  
 Db 1316 TATGACATGCGATATGCCAGCAACCTTCTGCTAAGAAAAGGACCAATGAAGGTTTC 1375  
 QY 401 HisGluAlaValGlyLysLysMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
 Db 1376 CATGAAGCTGTGGAGAAATCATGCTTCTGACATCAACCCCAAGCATCTGAATTC 1435  
 QY 421 TleGlyLeuLeuSerProAspPheGlnGlnAspAsnGlnLysTrpGluLysAsnPheLeu 440  
 Db 1436 ATTGCTCTTGGCATTCCGATTTTCAAGAAATGCGCAACAGAGATTAACCTTCACTG 1495  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrLysMetLeuGlnLysTrpArg 460  
 Db 1496 AAACAGGATTCGACAAATGTTGGAAACACTTACCTTACTTAACATGTAAGAAAGTGAGG 1555  
 QY 461 TrpMetValPheLysGlyLysLysLeuProLysAspLysTrpMetLysLysTrpPheGlnMet 480  
 Db 1556 TGGATGCTCTTCCGGGGTGAATATCCCAAGAGCAGAGATGAAGAAAGTGGAGATG 1615  
 QY 481 LysArgGluLysValGlyValGluProValProHisAspGluLysTrpCysAspPro 500  
 Db 1616 AAGCGGAGATCGTTGGTGTGTGGAGCTTCTGCTATGATGAACATCTGAGACCT 1675  
 QY 501 AlaSerLeuPheHisValSerAsnAspLysSerPheIleArgTrpTrpArgThrLeu 520

```

Db 1676 GCATCTCTTCATGTTCTTAATGATTAATCATTCATTCATGATNTATACAGAGACCATT 1735
QY 521 TTYGInPheGlnPheGlnGlnAlaLeuGlySGlnAlaAlaIleGlyLeuPheUHis 540
Db 1736 TACCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1795
QY 541 LysCysAspIleSerAsnSerThrGlnAlaGlnIleGlyLeuPheAsnMetLeuArgLeu 560
Db 1796 AATATGACATCTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1855
QY 561 GlyLysSerGlnPheTrpThrLeuAlaLeuGlnIleGlyLeuPheAsnMetLeuArgLeu 580
Db 1856 GGAATATTCAGAGCCCTGAGACCAAGCCCTTGAGAAATGCGTAGAGAGCAAGAAATGAT 1915
QY 581 ValArgProLeuLeuAsnThrPheGlnPheLeuPheUHisGlyLeuPheAsnMetLeuArgLeu 600
Db 1916 GTAAACACATCTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1975
QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTrpAlaAspGlnSerIleLeuVal 620
Db 1976 AATCTCTTTTGGGGTGGAGACATGATGAGCCCATATGCCACCAACCATTAATGATG 2035
QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTrpIleTrpAsnAspAsnGluMet 640
Db 2036 AGGATTAACCTTAATATGAGCTTGGAGCTAATGATGATGATGATGATGATGATGATGATG 2095
QY 641 TyrLeuPheArgSerSerValAlaTrpAlaMetArgIleTrpPheLeuLysValLysAsn 660
Db 2096 TTCCGTGTCGATCATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2155
QY 661 GlnMetIleLeuPheGlyGlnAlaAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 2156 CAGACAGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2215
QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleLeuProArgIleTrpVal 700
Db 2216 TTCTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2275
QY 701 GluLysAlaIleLeuMetSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2276 GAAATATCCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2335
QY 721 SerLeuGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2336 AGCTGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2395
QY 741 IleTrpLeuIleValPheGlyValAlaMetGlyValIleValGlyIleValIleLeu 760
Db 2396 ATATGCTGATTAATTTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2455
QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
Db 2456 ATTGTCTACTGATCAAAAGCTCGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCT 2515
QY 781 TyrAlaSerIleLeuSerIleGlyLysAsnAsnProGlyPheGlnAsnThrAsp 800
Db 2516 TATGACTCGATGACATTTGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGATGAT 2575
QY 801 ValGlnThrSerPhe 805
Db 2576 GCTCAGACTTCTT 2590

```

```

RESULT 18
LOCUS AB053181 2760 bp mRNA linear ROD 03-Apr-2001
DEFINITION Mus musculus ACE2 mRNA for angiotensin-converting enzyme-related
ACCESSION AB053181
VERSION AB053181.1 GI:13517092
KEYWORDS
SOURCE Mus musculus cDNA to mRNA.
ORGANISM Mus musculus

```

```

REFERENCE
AUTHORS Komatsu,T., Sugano,S., Imai,J., Suzuki,Y., Haneoka,K., Ymada,Y.,
Hida,M., Tanigami,A. and Muroi,S.
TITLE Molecular cloning, mRNA expression, and chromosomal localization of
mouse Angiotensin-converting Enzyme-Related Carboxypeptidase
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2760)
AUTHORS Komatsu,T., Sugano,S. and Suzuki,Y.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2001) Takami Komatsu, The institute of medical
science, University of Tokyo, Laboratory of Genome Structure
Analysis, 4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
(E-mail:komatsu@ims.u-tokyo.ac.jp, Tel:81-3-5449-5283(ex.75283),
Fax:81-3-5449-5416)
FEATURES
source
location/Qualifiers
1..2760
/organism="Mus musculus"
/db_xref="taxon:10090"
1..2760
/gene="ACE2"
187..2583
/gene="ACE2"
/codon_start=1
/product="angiotensin-converting enzyme-related
carboxypeptidase"
/protein_id="BAB40431.1"
/db_xref="GI:13517093"
/translation="MSSSSMLLSIVAVTASLSEENAKFFLNFEANDLSYSS
LASMNTNTTEPNNAKMSAAKSAFTEBDSKTAQSFSDQETPIKROLALQO
SGSSALSAKNNKOLLNTLNTSTYISGVCKPNPQSCLEPLEDEIMATSDYNS
KLWAGSGWAEEVKOLPLVEEYVYVKNMARANNYNDYCDWRDYEAEADQYNN
NLOLIDVEKTEFAEIKPLYEHLHAVRRKLMPTPSYISPTGCLPAHLGPMGRFMT
NLPLVPEPAOKFNIDVDAMNOCMDAERIPQEAKEFVSGVGLPHMTCGWMASMTL
EPADGRKVVCHPAMDLGSDPRIKMCTKRVMDNFTLAHHEMGTQYDMAVARPDL
RNGANGPEHVAEGLMSLSAATPKILKSTGILRSPQDSFEINFLIKALITVGLT
PEYMLEKKRWVRFGEITPEKMMKMKMEKREIYGVPELPHEDTYCDPSLTHVSN
DYSEITRYRTIYQFQFQALCOAKTKOSLHKCISNSTAGGLKMLSLGSEBW
TKALEYVVARMDVKPLNTYQPLEWLKEDQNRNFSYVGMTWSPYADOSIKVRIEL
KSLAGNAAYEMWNNEMFELRSSVAAAMRKYSIINKQYPLLEEDVRSYDLKPVSVY
FVYTSPOYNDVIPSEVEDAIRMSGRGINDFGINDNSLEFLGIHPLPEPPQPVY
ITWILINGVMAVVYGIILTYTGIKGKKKKNETKRESPPDSYGHMKRRKQCNIPRO
"
BASE COUNT 830 a 585 c 628 g 717 t
ORIGIN
Alignment Scores:
Pred. No.: 3,08e-290 Length: 2760
Score: 3562.00 Matches: 660
Percent Similarity: 89.45% Conservative: 61
Best Local Similarity: 81.89% Mismatches: 84
Query Match: 83.01% Gaps: 1
DB: 10
US-09-978-385-2 (1-805) x AB053181 (1-2760)
QY 1 MetSerSerSerSerThrPheLeuLeuLeuSerValAlaValThrAlaAlaGlnSerThr 20
Db 187 ATGTCCAGCTCTCTCTGCGCTCTCTCAGCCTTGTCGTTACTAGTCAGCTCCCTC 246
QY 21 IleGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGlnAspLeuPhe 40
Db 247 ACCGAGGAAATGCCCAAGACATTTTAAACACTTAATACGAGAAAGCTGACTGCT 306
QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTrpAsnIleThrGlnGlnAsnValGln 60
Db 307 TATCAAAAGTTCACTGCTCTTGGAAATTAATTAATCACTTACGAGAAAGAAAGCCCAA 366
QY 61 AsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuGlnGlnGlnSerThrLeuAla 80
Db 367 AAGATGAGTGAAGGCTGACGCCCAATGCTGCTTTTATGAAGAACAGCTTAAGACTGCC 426

```



Db 2587 TGCTCAGACTTCCTTT 2602

RESULT 19  
AX047765 2638 bp DNA linear PAT 15-Dec-2000

LOCUS AX047765  
DEFINITION Sequence 8 from Patent WO0070032.  
ACCESSION AX047765  
VERSION AX047765.1 GI:11876771

KEYWORDS  
SOURCE Mus sp.  
ORGANISM Mus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 2638)  
Piddington, C.S., Perrie, C.R., Shoemaker, K.E. and Bishop, P.D.  
Zace2: a human metalloenzyme  
Patent: WO 00/0032-A 8 23-NOV-2000;  
Zymogenetics, Inc. (US)

FEATURES  
source location/Qualifiers  
1..2638  
/organism="Mus sp."  
/db\_xref="taxon:10095"  
106..2523  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC18993.1"  
/db\_xref="GI:11876772"  
/translation="MSSSMILLIVAVTTAOSTEENAKTFINNTNOEADISYSS  
LASVNYNTTEENAKMSRAAKMSAFYEOKTAKOSELSLOETPIIKROLALQO  
SGSALADKNKOLNTLNTMSTTISGKVCANRNPOECLELPEGLDEIMASTYNS  
RUMABGHRAVQOLRPLEEVVVKNEKARNANNNDYGRGDEAGADGYNN  
NQIIEDEVERFAIKPLVLEHYVKNKIMDYPISTGCLPAHLGDMGRMT  
EPAGRRVYVCPAMWDGHDPRFKCTKVTMNFILAHKMGHIOYDMAVYROPPL  
RNGANECFHEAVGIMSLSATPKHLKSGILPSDPOEDSERINLLQALTTGTL  
PFTMLEKMMWVEFGEIPEKQMKMKMKKPRIVGVPEPLRDETYCDPASLFEVS  
DYSPRYRTITFOFOEALCOAKYNSLKCDISNTEAGQKLKMLSGNSEPW  
TEALENVYVGRNDYKPLRYFOPLDMLKEORNSFVWNTWESPAYDOQSIRVLSL  
KSLGANAIRYRTNMEFLKRSVAYAMRYSSITKNQVTFLEEDYRVSDELKRSFY  
PFTSPQNSVDVIRSEVEDAIRMSGRINDYFGINDNLEFLGIHPTLEPYOPPT  
TWLIRGVVALVVGVIILITVIGIKRRKKNETKREPNYSMDIGKESNAGRONS  
DDAQTSS"

BASE COUNT 797 a 557 c 615 g 669 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.55e-290 Length: 2638  
Score: 3561.00 Matches: 659  
Percent Similarity: 89.32% Conservative: 60  
Best Local Similarity: 81.86% Mismatches: 86  
Query Match: 82.99% Indels: 0  
Gaps: 0

US-09-978-385-2 (1-805) x AX047765 (1-2638)

QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20  
Db 106 ATGCCAGCTCCCTCCCTCCCTCCAGCTTGTCTTACTACTCTCAGTCCCTC 165  
QY 21 ILGLIUGLInalalysThrPheLeuAspLysPheAsnHisGluAlaGlnAspLeuPhe 40  
Db 166 ACCGAGAAATCCCAAGACATTTTAAACACTTTAATCAGAGGCTGAAGACCTGTCT 225  
QY 41 TyGlnSerSerLeuAlaSerTrpAsnTyAsnThrAsnIleThrGlnGlnAsnValGln 60  
Db 226 TATCAAAAGTTCACCTGCTGTTTGGAAATTTATTAATCACTTACTGGAAGAAATATCCCAA 285  
QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla 80  
Db 286 AAGATGAGTGTAGGCTCAGCCAAATGTCTGCTTTTATGAGAAACAGCTTAAGCTGCC 345  
QY 81 GlnMetYrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100  
||| ::: |||||||||||||||

Db 346 CAAAGTTTCACCTACAGAAATCCAGACTCCGATCATCAACCTCAACGACCTT 405  
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120  
Db 406 CACCAAGAGTGGGCTTCACACACTCTCAGCAGCAAGAACAAACAGTTGACCAATCTG 465  
QY 121 AsnThrMetSerThrIleYrSerThrGlyLysValLysAsnProAspAsnProGlnGlu 140  
Db 466 AACACCATGAGCCACCTTTACAGTCTGGAAAAGTTTCCAAACCCCAAGAACCCACAGAA 525  
QY 141 CysLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyAsnGlu 160  
Db 526 TGCTTATTACTTACGACCAAGATTTGATGAATAATGGGACACAGACAGATACAACTCT 585  
QY 161 ArgLeuThrAlaThrGlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuYr 180  
Db 586 AGGCTTGGGCATGGAGGGCTGGAGGCTGAGTTGGCAAGCAGCTGAGGCCCTGTAT 645  
QY 181 GlnGluTyValValLeuLysAsnGluMetAlaArgAlaAsnHisTyGluAspTyGly 200  
Db 646 GAAGAGTATGTGCTCTGTAACAGAGATGGCAGAGCAAGCAATTTAACGACTATGG 705  
QY 201 AspTyTrpArgGlyAspTyGluValAsnGlyValAspGlyTyAspTySerArgGly 220  
Db 706 GATTATTGGAGGGGAGCTATGAAACAGAGGAGCAGATGGCTCAACTATTAACCTTAAC 765  
QY 221 GlnLeuIleGluAspValGluHisThrPheGlnGluIleLysProLeuTyArgLysLeu 240  
Db 766 CAGTTGATGATGAAGATGTGAACGATACCTCCCAAGATCAACCATTTGATGAGCATCT 825  
QY 241 HisAlaTyValAlaGluAlaLysLeuMetAsnAlaTyProSerTyIleSerProIleGly 260  
Db 826 CATGCTCTTGGAGGAGGAAGTTGATGATACCTTCCATCAATCACCCACCTAC 885  
QY 261 CysLeuProAlaHisLeuLeuGluIleAspMetTrpGlyArgPheThrAsnLeuTySer 280  
Db 886 TGCTCCCTGCCCATTTGCTTGATATGATGGGGTATGATTTGGACAAATCTGTACCT 945  
QY 281 LeuThrValIleProPheGlnGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
Db 946 TTGACTGTTCCTTGGCAGCAAAACCAACATGATGATCAATGATGATGATGATGATGATGAT 1005  
QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320  
Db 1006 GGCTGGATGCGAAGAAAGATATTTCAAGAGCGCAGAAATTTCTTCTGTCTGTGGCTT 1065  
QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340  
Db 1066 CCTCATATGACTCAAGGATTTCTGGCAAACTCTATGCTGACTGACGACCAAGATGGCGG 1125  
QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
Db 1126 AAAGTTGTCTGCCACCCCAAGCTTGGGATCTGGGACAGGAGACTTCCGAATCAAGATG 1185  
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetAlaHisIleGln 380  
Db 1186 TGTCAAAGGTTCACAAATGAGCACAATCTTGACAGCCCATCAGAAATGAGACATCCAA 1245  
QY 381 TyrAspMetLeuAlaTyAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyPhe 400  
Db 1246 TATGACATGCGCATATGCGCAGCAAGCAACTTCTGCTCAAGAAAGGAGCCCAATGAAGGTTTC 1305  
QY 401 HisGluAlaValGlnIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
Db 1306 CATGAAGCTGTGAGAAATCATGTCATTTCTGACAGCTACCCCAAGCATGTGAATATCC 1365  
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440  
Db 1366 ATTTGCTTCTGCGCATCCGATTTTCAAGAAATAGCAAAACAGAGATTAACCTTCTACTG 1425  
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyIleMetLeuGlnLysTrpArg 460  
Db 1426 AACACGCAATGACATTTTGGAAACACTACCGTTTACTTACATGTTTAGAAGATGGAGG 1485



```

Db      541 GARGARTACTNGTNTYNNARAAYGARATGGCGMGGCAATCATATGARGATATYGCN 600
Qy      201 ASPTYTRPARGGLYASPTRYGLUVALASNGLYVALASPGLYTYRASPTRYSEARGLY 220
Db      601 GATTATGGMGNGNGNAYARAGRNATYAGGNGTNGAYGGNTAYGATTAYWSMNGN 660
Qy      221 GlnLeuIleGluAspValGluHisThrPheGluGluIleuysProLeuTyrGlnHisLeu 240
Db      661 CARYTNATGARGAYGTNCARACACNTTTCARGARATHAARCCNTTATGARGCAAYTN 720
Qy      241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db      721 CAGCNTATGTMGNCNNAARNTATGAGCNCNTATCCNWSNTATATWSMCCNATHG 780
Qy      261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpHisLeuTyrSer 280
Db      781 TGYTNCCGNCNCAYTNTTNGNCATATGTTGGGNGMNTTTCGACNAAAYTNTAYWSN 840
Qy      281 LeuThrValProPheGlyGluLysProAsnIleAspValIleHisAspAlaMetValAspGln 300
Db      841 YTNACGTCCNTTTCGNCARARACNNAAYATHGATGNACNGAYGCNATGTCATCAR 900
Qy      301 AlaTrpAspAlaGluArgIlePheGlyGluAlaGlyLysPhePheValSerValGlyLeu 320
Db      901 GCGTGGATGCGNCARMCNATHTTAAARGARCCNARARATTTTTCNWSNGTNGNVTN 960
Qy      321 ProAsnMetTrpGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
Db      961 CCNAAATGACNCARGNTTTCGGARAAYSNATGTTMNCNGAYCCNCGNAAATGTCAR 1020
Qy      341 LysAlaValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db      1021 AARGNGTNTGTCACCCNACGNCNTGGAYTTNGNNAARGNGAYTTTCNATHTYTNATG 1080
Qy      361 CysThrLysValIleHisAspAspPheLeuThrAlaHisGlyMetGlyHisIleGln 380
Db      1081 TGYACNNAARGTNACNATGATGATTTTTCNACNCCNATCAYGARGATGGNCAYATHCAR 1140
Qy      381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyPhe 400
Db      1141 TAYGATATGGCNTAYGCGNCNARCCNTTTCNATNTNMGNAAAGNCGNNAAYGARGCNTTY 1200
Qy      401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db      1201 CAYARGCNGTNGNARGATHTATGWSNTTMSMCCNCCNACNARCAAYTNTAARWSN 1260
Qy      421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
Db      1261 ATHGAGTNTNTMNSCCNCAATTCARGARGATAAAGARCCNARATHTAATTTTNTNTN 1320
Qy      441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
Db      1321 AARCARCNGTNTNACNATHGNGNACNYTNCNTTTCNATNTATGTYNGARARATG 1380
Qy      461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpGluMet 480
Db      1381 TGGATGCTNTTTCARCGNARGARATHCCNAAARGATGATGATGAARARATGGTGGARRAG 1440
Qy      481 LysArgGluIleValGlyValIleGluProValProHisAspGluThrTyrCysAspPro 500
Db      1441 AARGNARGARATHGNGNNGTNGARCCNGTNCNCAATGARGARCCNATYTGCGYCCN 1500
Qy      501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
Db      1501 GCNMSNTTNTTCAYGTMNSNAAAYGATYATWSNTTATATHTMGNTATATYACNMGNACNNTN 1560
Qy      521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHis 540
Db      1561 TAYCARTTCARTTCARARGRCNTTNTTCARGCNCNAAARCAATGARGCNCNTTTCAY 1620
Qy      541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560

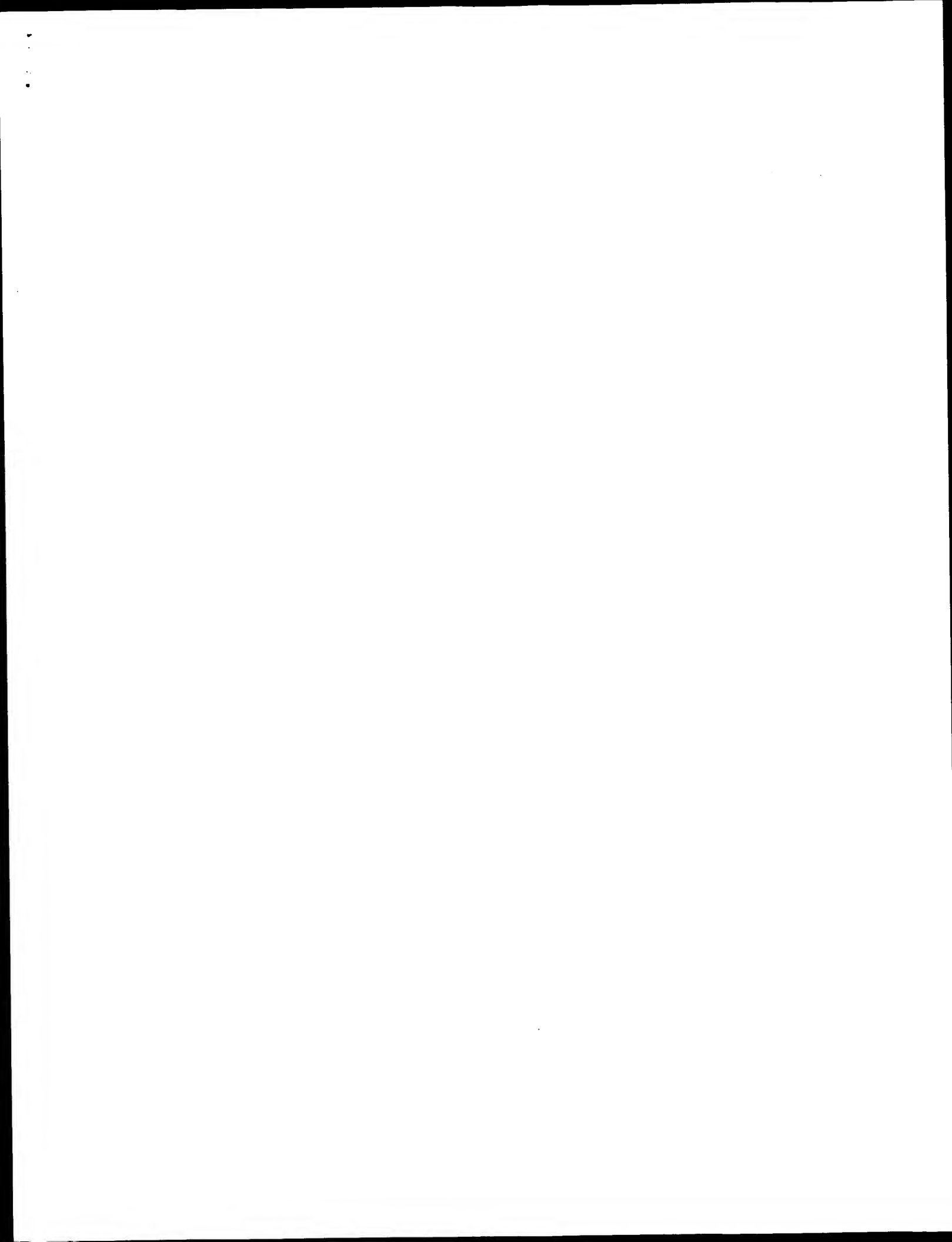
```

```

Db      1621 AARTGCAVATHTWSNAAYSNACNGARCGCNCARARNTTNTAAATGYTNMGNTN 1680
Qy      561 GlyLysSerGluProTrpThrLeuAlaGluAsnValValGlyAlaLysAsnMetAsn 580
Db      1681 GGNAAARWSNARGCNCNTGNCNATNTGNTNGARAAATGNTGNGCNAARAAATGAA 1740
Qy      581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpPheLysAspGlnAsnLys 600
Db      1741 GTMNCNCNTTTCNAAATYATTCARGCNCNTTNTTCNTGNTGNTNAAARGATCAARAAAR 1800
Qy      601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
Db      1801 AAATWSNTTTCGNGTNGGNSNACNAYTGWSNCCNATYCGNCAATGARGWSNATHAARGTN 1860
Qy      621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640
Db      1861 MGNATHTWSNTTAAARWSNCCNTNGNCNAYARCCNTATGARGTAATGARGATG 1920
Qy      641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660
Db      1921 TAYTNTTTCGWSMWSMWSMNGTNCNATYGCNATGMCNARFATYTTTNTAARGTNAARAY 1980
Qy      661 GlnMetIleLeuPheGlyGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db      1981 CARATGATHTTTCGNGARGARATGNTMNGTNGCNAAATYTNAAARCCNMGNATHTWSN 2040
Qy      681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTrpGluVal 700
Db      2041 TTYAATTTTTCGTCNACGNCNCAARAYGTNMSNATATHTHCCNMGNCNARGTNT 2100
Qy      701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db      2101 GARAARGCNATHMGNATGWSMWSMNGNATHAAYGAYCNCNTTTCNATNTAAYGATYAY 2160
Qy      721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db      2161 WSNYTNAGARTTTCGNGNATHCARCCNACNVTNGNCCNCAATYCARCCNCCNNTMWSN 2220
Qy      741 IleThrPheIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeu 760
Db      2221 ATHGCTTNTATHTGNTTTCGNGTNGTATGNGTNTATHTGNTGNTGNTATHTGNTATHTN 2280
Qy      761 IlePheThrGlyIleArgAspArgLysLysAsnLysAlaArgSerGlyGluAsnPro 780
Db      2281 ATHTTTCNCGNATHTMGNGAYMGNAARAAARAAARAAARCCNMGNSNGNGARRAAYCCN 2340
Qy      781 TyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAsp 800
Db      2341 TAYGCNMSNATHTGAYATHTWSNAAARGNGARAAVAAVCCNCGNTTTCARAAAYCNCAAY 2400
Qy      801 ValGlnThrSerPhe 805
Db      2401 GTNCARCCNMSNTTY 2415

```

Search completed: March 7, 2003, 09:08:57  
 Job time : 4946 secs





Run on: March 7, 2003, 06:33:34 ; Search time 406 Seconds

File: US-09-978-385-2  
Perfect score: 4291

xgapop 10.0 , xgapext 0.5

Searched: 2185239 scans 1125000150

Minimum DB seq length: 0

Maximum Match 100%

Command line parameters:

```

nucleotide+frame+g2n.model -DEV=xip
0-/cgn2.1/USPRO.spool/US09978385.r/unat.28022500.104716.1332/app.query.fasta.1.967
-DB=N.Genesec.101002 -OMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsim2 -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=20
-MODE=LOCAL -OUTFILE=plo -NORM=ext -HAPSIZES=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09978385.@CGN.1.1396.@runat.28022500.104716.1332 -NCPU=6 -ICPU=3
-NO.WALYPX -NO.MMAP -FAREQ=1 -NG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-XLRP.TIMEOUT=30 -THEADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database : N\_Geneseq\_101002:\*

1:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1980.DAT *
2:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1981.DAT *
3:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1982.DAT *
4:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1983.DAT *
5:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1984.DAT *
6:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1985.DAT *
7:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1986.DAT *
8:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1987.DAT *
9:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1988.DAT *
10:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1989.DAT *
11:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1990.DAT *
12:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1991.DAT *
13:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1992.DAT *
14:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1993.DAT *
15:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1994.DAT *
16:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1995.DAT *
17:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1996.DAT *
18:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1997.DAT *
19:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1998.DAT *
20:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA1999.DAT *
21:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA2000.DAT *
22:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA2001A.DAT *
23:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA2001B.DAT *
24:	/SID52/gcgdata./geneseq./geneseqn -emb1./NA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	4291	100.0	2418	21	AA59465	Human MPROT15 cod
2	4291	100.0	3334	22	AA584366	Human Zace2 prote
3	4291	100.0	3396	21	AA917664	cdna encoding a h
4	4291	100.0	3396	22	AAAD07258	Human angiotensin
5	4291	100.0	3396	24	ABK87623	cdna encoding hum
6	4291	100.0	3396	24	AAAD32586	Human ACE-2 full-
7	4142	96.5	3732	22	AAAS21279	Human CDNA sequen
8	4061	94.6	2920	22	AAAS14880	Human CDNA encodi
9	4013	93.5	2911	22	AAAS14890	Human CDNA encodi
10	3740.5	87.2	2652	21	AA545466	Human CDNA encodi
11	3579	83.4	2638	22	AAAC8368	Human MPROT15 cod
12	3561	83.0	2638	22	AAAC8370	Mouse Zace2-5 pro
13	3509	81.8	2638	22	AAAC8433	Mouse Zace2-10 pro
14	3291	76.7	2350	22	AAAC84367	Human G-protein-c
15	3119	72.7	3474	22	AAAS42515	Human CDNA encodi
16	2899	67.6	2415	22	AAAC43659	Mouse Zace2-5 pro
17	1344	31.1	2477	12	AAAO10338	Human Zace2 prote
18	1337	31.2	4020	21	AAAS38330	Human CDNA encodi
19	1337	31.2	4022	24	AAK93395	Human angiotensin
20	1337	31.2	4024	11	AAQ04037	DNA of APP relatee
21	1337	31.2	4024	10	AAAX58580	Human angiotensin
22	1336	31.1	4020	19	AAVA13320	Human angiotensin
23	1334	31.1	3939	22	AAAS6085	Angiotensin convers
24	1334	31.1	4563	22	AAAS6057	Angiotensin conver
25	1310	30.5	3942	20	AAAX58581	Rat angiotensin ce
26	1275	29.7	5005	22	AAH57430	Human intestin ce
27	1240.5	28.9	2568	24	AAK99739	CDNA encoding the
28	1066	25.3	2089	23	ABLI4379	Drosophila melanog
29	1057	24.6	2074	16	AAO82948	Tick carboxypeptid
30	1028	24.0	2450	23	ABLI6657	Drosophila melanog
31	961	22.4	9006	22	AAH78783	Nucleotide sequenc
32	941.5	21.9	5632	23	ABLI4478	Drosophila melanog
33	919.5	21.4	5060	23	ABLI6696	Drosophila melanog
34	841	19.6	2379	24	ABU28623	Drosophila melanog
35	769	17.9	561	24	ABQ57861	Human colon cancer
36	721	16.8	2082	21	AAAS46692	Degenerate sequenc
37	721	16.8	2082	22	AAAS14169	Human zinc metallo
38	721	16.8	2082	22	ABU40399	Degenerate coding
39	715.5	16.7	1395	22	AAH78785	Nucleotide sequenc
40	597.5	15.0	446	24	ABQ57935	Human colon cancer
41	587.5	13.9	42573	23	ABU28622	Drosophila melanog
42	502.5	11.7	2025	23	ABU04671	Drosophila melanog
43	476	11.1	313	20	AAAV8528	EST clone AU47. H
44	476	11.1	1836	23	ABU27143	Drosophila melanog
45	469	10.9	467	19	AAV09277	Nucleotide sequenc



Db 1621 AATGTCACATCTCAAACTGTACAGAACTGGACAGAACTGTTCAATATGCTGAGGCTT 1680  
 QY 561 G1YLSerSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetLsn 580  
 Db 1681 GGAATAATCAGAACCTTGACCTGACCTAGCATGTGAAATATGTTGTAGAGCAAGAACATGAAT 1740  
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600  
 Db 1741 GTAAGGCCACCTGCTCACTACTTGTGACCCCTTATTTACTGCTGAAAGACAGAACAAAG 1800  
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTrpThrLaspGlnSerIleLysVal 620  
 Db 1801 AATTCCTTTGGGATGGAGTACCGACTGAGTCCATATGCAACCAACCAATCAAGTG 1860  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGlnTrpPaspAsnAspGlnMet 640  
 Db 1861 AGGATAAGCCTAAATACGCTCTTGGAGATAAAGCATATATGAAAGCAATGAAATG 1920  
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgIleTrpPheLeuLysValLysAsn 660  
 Db 1921 TACCTGTTCCGATCATCTGTTGCATATGATGAGCAGTACTTTTAAAGTAAAAAAT 1980  
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 Db 1981 CAGATGATCTCTTTGGGAGAGAGATGCGAGTGCAGTGCATATTTGAAACCAACAACTCC 2040  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTrpGluVal 700  
 Db 2041 TTTAATTTCTTTGTGACATGCACTAAATAATGTCTGATTCATCTTCTAGAACTGAACTT 2100  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 Db 2101 GAAAGGCCATCAGATGTCGCCGAGCCGATCATGATGCTTCCCTGATGACAAAC 2160  
 QY 721 SerLeuLuhPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 Db 2161 AGCCTAAGTTCTGGGGAGAACACCAACCTGGACCTCTAAACCAACGCCCTGTTTCC 2220  
 QY 741 IleTrpLeuIleValPheGlyValIleMetGlyValIleValGlyIleValIleLeu 760  
 Db 2221 AATAGGCTGANTGTTTGTGGAGTTGTGATGGAGTGTATGTGTGTCATTCATCTG 2280  
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro 780  
 Db 2281 ATCTTCACCTGGGATCAGAGATCGGAAAGAAAATTAACCAAGTGGAGAAAATCCT 2340  
 QY 781 TyrIleSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAsp 800  
 Db 2341 TATCCCTCCATCGATATTAGCAAGAGAAAATATATCCAGATTCCAAAACACTGATGAT 2400  
 QY 801 ValGlnThrSerPhe 805  
 Db 2401 GTTCAGACCTCTTT 2415  
 RESULT 2  
 AAC84366  
 ID AAC84366 standard; cDNA; 3334 BP.  
 XX  
 AC AAC84366;  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 DE Human Zace2 protein encoding cDNA.  
 XX  
 KM Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;  
 KM zinc metalloproteinase; blood pressure; zinc protease; hypertension;  
 KM ventricular systolic dysfunction; renal impairment; heart failure;  
 KM scleroderma renal crisis; atherosclerosis; . antiinflammatory; human;  
 KM antiarthritic; bradykinin inactivator; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers

FT CDS 35..2452  
 FT /\*tag= a  
 FT /product= "Zace2"  
 PN MO200070032-A1.  
 XX  
 XX 23-NOV-2000.  
 XX  
 XX 03-MAY-2000; 2000MO-US11932.  
 XX  
 PR 13-MAY-1999; 99US-0311482.  
 PR 27-AUG-1999; 99US-0384706.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;  
 DR WPI, 2001-025018/03.  
 DR P-PSDB; AAB48095.  
 XX  
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory  
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases  
 PT associated with inflammation such as arthritis and enterocolitis -  
 XX  
 XX Example 1; Page 95-100; 125pp; English.  
 CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-  
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood  
 CC pressure regulation and fertility. Zace2 can be expressed by standard  
 CC recombinant methodology. Zace2 polypeptides are useful for treating an  
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),  
 CC diseases associated with inflammation like arthritis and enterocolitis,  
 CC as targets for identifying modulators of zinc protease activity, for  
 CC screening or identifying new angiotensin-converting enzyme (ACE)  
 CC inhibitors, and as a basis for rational drug design for inhibitory  
 CC molecules. The nucleic acids can be used to detect the expression of a  
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and  
 CC for detecting and localizing Zace2 gene expression in tissue samples,  
 CC to determine whether a subject's chromosomes contain a mutation in the  
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.  
 CC Inhibitors of ACE are used for treating hypertension of various  
 CC conditions, including left ventricular systolic dysfunction, progressive  
 CC renal impairment, scleroderma renal crisis, congestive heart failure due  
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be  
 CC used to treat infertility while Zace2 antagonists are used for inducing  
 CC infertility. The present sequence represents a cDNA encoding the human  
 CC Zace2 protein.  
 CC  
 SQ Sequence 3334 BP; 1011 A; 640 C; 754 G; 929 T; 0 other;  
 XX  
 XX  
 Alignment Scores:  
 Pred. No.: 0 Length: 3334  
 Score: 4291.00 Matches: 805  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-978-385-2 (1-805) x AAC84366 (1-3334)  
 QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20  
 Db 35 ATGTCAAGCTCTTCCTGGCTCTTCACGCTTGTGTGTAACGTGCTGACGTCACAC 94  
 QY 21 IleGluGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40  
 Db 95 ATTAGAGCAACGGCCAGACATTTTGGACAGATTAAACAGAGCCGAGACCTGTTC 154  
 QY 41 TYGILSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGln 60  
 Db 155 TATCAAAAGTCACTTCTTGTGAATTTAAACACCAATATATACGAGAGAAATGCCAA 214  
 QY 61 AsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80

```

Db 215 AACATGAATATGCTGGGGCAATGCTGCTTTTAAAGAAACAGTCCACACTTGGC 274
QY 81 GlnMetTyrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnIleGlnAlaLeu 100
Db 275 CAATGTATCCACTACAGAAATTCAGAAATTCACAGTCAAGCTTACGTGACGGCTCTT 334
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
Db 335 CAGCAAAATGGGTCTTCAGTGGCTTCAGAAAGCAACAGAGCAAGCGTTGAACAAATTCGA 394
QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 395 AATACATAGTCACACATCTACAGTACTGAAAAAGTTGTAAACCCAGATATCCACAAGAA 454
QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 455 TCGTTTATCTTGAACCAAGGTTTGAATGAATTAAGCAACAGTTTAACTACATCAATGAG 514
QY 161 ArgLeuThrAlaIleProLysSerThrArgSerGluValGlyGlnLeuArgProLeuTyr 180
Db 515 AGGCTGTGGGCTGGGAAAGCTGGAGATCTGAGTGGGAGCAAGCTGAGCGCATTAAT 574
QY 181 GluGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
Db 575 GAAGAGTATGTGCTCTTGAATAATGAGATGCAAGAGCAAAATCATTAATGAGACTATGGG 634
QY 201 AspTyrThrArgGluAspTyrGluValAsnGluValAspGlyTyrAspTyrSerArgGly 220
Db 635 GATATATGGAGAGAGACATATGAAATGGGATGATGGCTATGACTACAGCGCGGC 694
QY 221 GlnLeuIleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlnHisLeu 240
Db 695 CAGTTGATGAAAGATGTGGAACATACCTTTGAAGAGATTAACCATTAATGACATCTT 754
QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 755 CAGGCTATGTGAGGGCAAGTTGATGAAATGCTATTCCTATATCATGCTCAATGTGA 814
QY 261 CysLeuProAlaHisLeuLeuGlnLysAspMetTyrGlyArgPheThrThrAsnLeuTyrSer 280
Db 815 TGGCTCCCTGCTCATTTGCTTGTGTATATGTTGGGATGAGTTTGGACAAATCTGTACTCT 874
QY 281 LeuThrValProPheGlnLysGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 875 TTGACAGTTCCTCTTGGACAGAAACCAACATAGATGATGATGATGATGATGATGATGAT 934
QY 301 AlaTyrAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerValGlyLeu 320
Db 935 GCCTGGATGACAGAGATATTCAGAGGCGGAGAGAGTCTTGTATCTGTTGCTCTT 994
QY 321 ProAsnMetThrGlnGlyPheThrPgluAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 995 CCTAATATGACTCAAGAGATTTCTGGAAAAATTCATGCTACACGACCCAGGAATGTTGAG 1054
QY 341 LysAlaValLysHisProThrAlaThrAspLeuGlnLysGlyAspPheArgIleLeuMet 360
Db 1055 AAAGCAGCTCGCCATCCACAGCTTGGGAGCTGGGGAAGGGGAGCTTCAAGATCTTATG 1114
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHisIleGln 380
Db 1115 TGCACAAAGTACAAATGAGACACTTCCGACAGCTCATCAAGAAATGGGGATATTCAG 1174
QY 381 TyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyAlaAsnGlnLysPhe 400
Db 1175 TATGATATGCAATATGCTGCACAAACCTTTTCTGCTAAGAAATGAGCTTAATGAAGATTC 1234
QY 401 HisGluAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1235 CATGAAGCTGTTGGGAAATCATGTCACTTCTGACAGCACACCTTAACCATTTAAATCC 1294
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440

```

---

```

Db 1295 ATTGCTCTCTGTGACCCGATTTTCAAGAAACATGAACAGAAATAACTTCTGCTC 1354
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTyrArg 460
Db 1355 AAACAGACAGCTACGATGTTGGAGCTCTGCAATTTACTTAACATGTTAAGAAATGGAGG 1414
QY 461 TrpMetValPheLysGlyLysIleProLysAspGlnIlePheLysLysTyrPheGluMet 480
Db 1415 TGGATGGTCTTTAAAGGGGAATTTCCAAACACAGTGGATGAAAGAAAGTGGTGGAGATG 1474
QY 481 LysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCysAspPro 500
Db 1475 AAGCGAGATAGTGGGGTGGTGGAACTTGGCCCATGATGAACATATCTGAGACCCC 1534
QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
Db 1535 GCATCTGTTCCATGTTTCTTAATGATTAATCAATTCATTCATGATTAACACAGACCCCT 1594
QY 521 TyrGlnPheGlnPheGlnGlnAlaLeuGlnGlnAlaLysHisGlnGlyProLeuHis 540
Db 1595 TACCAATTCAGTTCAGAAAGCACTTGTCAACACACTAAACATGAAGGCCCTCTGCAC 1654
QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPheAsnMetLeuArgLeu 560
Db 1655 AAATGTGACATCTCAAACTCTACAGAAAGCTGGACAGAAAGCTGTTCAATATGCTGAGCCT 1714
QY 561 GlyLysSerGluProThrPheLeuAlaLeuGlnLysAsnValGlyAlaLysAsnMetAsn 580
Db 1715 GGAATATCAGAAACCTTGACCCCTAGCAATGGAAATTTGTAGAGCAAAAGAACATGAAAT 1774
QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPheLysAspGlnAsnLys 600
Db 1775 GTAAGGCACATGCTCACTACTTGTGACCCCTATTATTCGTGGCGTGAAGAACAGAACAG 1834
QY 601 AsnSerPheValGlyTyrPheThrAspTyrPheProTyrAlaAspGlnSerIleLysVal 620
Db 1835 AATTCTTTTGGGATGAGTACCGACTGGAGTCCATATGCAGACCCAAACATCAAAAGTG 1894
QY 621 ArgIleSerLeuLysSerAlaLeuGluAspLysValTyrGluThrPheAsnAspAsnGluMet 640
Db 1895 AGGATTAAGCTTAATATACCTCTTGGAGATTAACATATGAATGAAGCAATGAATGAAATG 1954
QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660
Db 1955 TACTGTTCCGATCATCTGTTGCTATGCTATGAGCAGACTTTTAAAGTAAAT 2014
QY 661 GlnMetIleLeuPheGlnGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 2015 CAGATGATTTCTTTTGGGAGAGAGATGTGCGAGTGGCTTAATTTGAACCAAGAAATCTCC 2074
QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
Db 2075 TTTAATTTCTTGTACCTGACCTTAATAAATGTGCTGATATTCATTCCTGAACCTGAAGTT 2134
QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2135 GAAAGGCCAATCAGAGATGCCCGAGCGATATCATATGCTTCCGTGATGATGACAAAC 2194
QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2195 AGCCTAGAGTTTCTGGGGATACAGCCAAACACTTGGACCTCTTAACCAAGCCCTGTTTCC 2254
QY 741 IleThrLeuIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeu 760
Db 2255 ATATGCTGATTTGTTTGGAGTTTGATGGAGTGTATGATGCTATGCTATGCTATGCTG 2314
QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
Db 2315 ATCTTCACTGGATCAAGATTCGGAAGAAAGAAATTAAGCAAGATGAGAAATATCCT 2374
QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800
Db 2375 TATGCCCTCATCATATTTACAAAGAGAAATTAATCCAGAGATTCACAAAACACTGATGAT 2434

```

OY 801 ValGlnThrSerPhe 805  
 DB 2435 GTTCAGACCTCCTTT 2449  
 RESULT 3  
 ID AAA12764  
 AA12764 standard; cDNA: 3396 BP.  
 AC AAA12764;  
 XX  
 DT 25-JUL-2000 (first entry)  
 DE cDNA encoding a human angiotensin converting enzyme-2 (ACE-2).  
 XX  
 KW Human: angiotensin converting enzyme-2; ACE-2; angiotensin I; Ang. (1-9);  
 KW blood pressure; hypertension; congestive heart failure; atherosclerosis;  
 KW chronic heart failure; acute heart failure; myocardial infarction;  
 KW renal failure; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 82..2499  
 FT /tag= a  
 FT /product= "angiotensin converting enzyme-2"  
 FT 82..135  
 FT sig\_peptide  
 FT /\*tag= b  
 XX  
 FN W0200018899-A2.  
 XX  
 PD 06-APR-2000.  
 XX  
 PF 29-SEP-1999; 99WC-US22976.  
 XX  
 PR 30-SEP-1998; 98US-0163648.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Acton LS, Robison KE, Hsieh FY;  
 XX  
 DR WPI: 2000-293140/25.  
 DR P-RSDB; AA184562.  
 XX  
 PT Isolated nucleic acid encoding angiotensin converting enzyme-2 (ACE-2)  
 PT polypeptide useful for detecting an ACE-2 therapeutic for treating  
 PT hypertension, congestive heart failure, myocardial infarction,  
 PT atherosclerosis and renal failure -  
 XX  
 PS Claim 1; Fig 1; 138bp; English.  
 XX  
 CC The present sequence encodes a human angiotensin converting enzyme-2  
 CC (ACE-2). ACE-2 is expressed predominantly in kidneys and testis. The  
 CC sequence of the full length ACE-2 cDNA was determined from a clone  
 CC obtained from a cDNA library prepared from mRNA of a human heart of  
 CC a subject who had congestive heart failure. ACE-2 has significant  
 CC sequence homologies with ACE enzymes, and has also been shown to  
 CC hydrolyse angiotensin I into Ang. (1-9). The ACE-2 therapeutics are  
 CC used to treat blood pressure related diseases and conditions, such as  
 CC hypertension, congestive heart failure, chronic heart failure, acute  
 CC heart failure, myocardial infarction, atherosclerosis and renal  
 CC failure.  
 XX  
 SQ Sequence 3396 BP; 1034 A; 660 C; 771 G; 931 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0  
 Score: 4291.00 Length: 3396  
 Percent Similarity: 100.00% Matches: 805  
 Best Local Similarity: 100.00% Conservatives: 0  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0  
 DB: 21

US-09-978-385-2 (1-805) x AAA12764 (1-3396)  
 OY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaIleGlnSerThr 20  
 DB 82 ATGTCAGACTCTTCCTGCTGCTCTTCAGGCTTGTGCTGTAACCTGCTCAGTCAC 141  
 OY 21 IleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaLysPhe 40  
 DB 142 ATTGAGGACAGCCCAAGACATTTTGGACAAAGTTTAAACACGAAACCCAGACCTGTC 201  
 OY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGlnValGln 60  
 DB 202 TATCAAAAGTTCACCTTCTTGGAAATTATACACCAATTTACGAGAGATGTC 261  
 OY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuGluGlnSerThrLeuAla 80  
 DB 262 AACATGAATTAATCTGGGACAAATGCTGCTTTTAAAGAAACAGTCACACCTGTC 321  
 OY 81 GlnMetTyrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100  
 DB 322 CAATGATTCACCTACAGAAATTCAGAAATTCACAGCTTCAGCTGAGGCTCTT 381  
 OY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120  
 DB 382 CAGCAAAATGGCTCTTCAGTGCCTCAGAGACAGACAAACGGTTGAAACCAATTCTA 441  
 OY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140  
 DB 442 AATACATGATGAGCCACATCTACAGTACTGAGAAAGTTTGAACCCAGATTAACCAAGAA 501  
 OY 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGlu 160  
 DB 502 TGCCTTATTCTTGAAACAGAGTTTGAATGAATGAATGAACAGTTTAACTACATGAG 561  
 OY 161 ArgLeuTrpAlaTrpLysSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180  
 DB 562 AGGCTCTGGGCTGGGAAAGCTGAGAGATCTGAGGTGGCAACCACTGAGGCAATTATAT 621  
 OY 181 GlnGluTyrValValLeuLysAsnGluMetAlaAlaGlnHisIleTyrGluAspTyrGly 200  
 DB 622 GAAGAGTATGAGTCTTGAATAAATGAGATGAGAGCAAGCAAACTATTAGAGCATATGAG 681  
 OY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220  
 DB 682 GATTATGAGAGAGAGACATGATGAATATGGGCTGATGATGATGATGATGATGATGATGAT 741  
 OY 221 GlnLeuIleGluAspValGlnHisThrPheGluGlnIleLysProLeuTyrGluHisLeu 240  
 DB 742 CAGTTGATTTGAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 801  
 OY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260  
 DB 802 CAGGCTATGATGAGGCAAAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 861  
 OY 261 CysLeuProAlaHisLeuLeuGlnLysPheMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280  
 DB 862 TGCCTCCCGCAGCATTTGCTGCTGATATGATGATGATGATGATGATGATGATGATGATGAT 921  
 OY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 DB 922 TTGACAGTTCCTTGGACAGAAACCAACATAGATGTTACTGATGATGATGATGATGATGAT 981  
 OY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320  
 DB 982 GCCTGGATGACAGAGATATATCAAGGAGGCGAGAGATTTTGTGTGTGCTTT 1041  
 OY 321 ProAsnMetThrGlnGlyPheTrpLysAsnSerMetLeuThrAspProGlyAsnValGln 340  
 DB 1042 CCTAATATGATCAACAGATATTCGGGAAATTCATGCAACGAGCAACGAAATGTTCTAG 1101  
 OY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 DB 1102 AAAGCAGTCTGCAATCCACAGCTTGGAGACCTGGGGAAGGCGAGCTTCAGATCTTATG 1161

QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisHISGLMetGlyHisIleGln 380  
 |||||||  
 Db 1162 TGCACAAAGGTGACATGGACACTTCTGCAGAGCTCATGAGATGGGCATATCCAG 1221  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluPhe 400  
 |||||||  
 Db 1222 TATGATATGGCATATGCGACAAACCTTTCTGTAAGAAATGGAGTAATGAAGATTC 1281  
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
 |||||||  
 Db 1282 CATGAAGCTGTGGGAAATCATGTCACTTCTGCAGCCACACCTTAACATTAAATCC 1341  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440  
 |||||||  
 Db 1342 ATTGCTCTTCTGCACCCGATTTTCAAGACAAATGAAACGAAATTAACCTCTGCTC 1401  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTyrArg 460  
 |||||||  
 Db 1402 AAACAAGCAGCTCAGATTGTGGAGCTGCGCACTTACTTACATGTAGAGAGTGAGG 1461  
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTyrTrpGluMet 480  
 |||||||  
 Db 1462 TGGATGGCTTTAAAGGGGAAATCCCAAGACAGTGGATGAAAAGTGGGGAGATG 1521  
 QY 481 LysArgGluIleValGlyValGluProValProHisAspGluThrTyrCysAspPro 500  
 |||||||  
 Db 1522 AAGCGAGAGATAGTGGGGTGGTGGAGACTGTGCCCTCATGAAACATACGTGATCC 1581  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520  
 |||||||  
 Db 1582 GCATCTCTGTCATGTTCTTAATGATTAATCATTCATTCGATTCGATTAACAAGACCTT 1641  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlyAlaAlaLysHisGluGlyProLeuHis 540  
 |||||||  
 Db 1642 TACCAATTCAGATTTCAAAGAACACTTGTCAAGCAGCTAAACAGAAAGCCCTGAC 1701  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyIleLysLeuPheAsnMetLeuArgLeu 560  
 |||||||  
 Db 1702 AATGTGACATCTCAAACTCTACAGAGCTGACAGAAACCTTCAATATGTGTAGGCTT 1761  
 QY 561 GlyLysSerGluProTyrPheLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsn 580  
 |||||||  
 Db 1762 GGAAATATCAGAACCTGACCTAGCATGTGAAATGTGTGAGGACAAAGACATGAAT 1821  
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspLysAsnLys 600  
 |||||||  
 Db 1822 GTAAAGCCACTGCTCAACTACTTGAAGCCCTTATTACTGCTGTAAGACCAAGACAG 1881  
 QY 601 AsnSerPheValGlyTyrSerThrAspTyrSerProTyrAlaAspGlnSerIleLysVal 620  
 |||||||  
 Db 1882 AATTTCTTTGTGGATGGAGTACCGACTGAGTCCATATGCAAGACCAAGATCAAAATG 1941  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGlyTrpAsnAspAsnGluMet 640  
 |||||||  
 Db 1942 AGGATTAAGCCTAAATCACTCTTGGAGATAAGCATATGAATGAAGACAAATGAATG 2001  
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660  
 |||||||  
 Db 2002 TACCTGTTCGATCATCTGTGCATATGCTATGAGGCACTACTTTTAAAGTAATAAAT 2061  
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 |||||||  
 Db 2062 CAGATGATCTTTTGGGAGAGGAGATGCGAGTGGCAATTTGAACCAAGATTCGCC 2121  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700  
 |||||||  
 Db 2122 TTTAAATTTCTGTCTCACTCCACTTAATAATGTGTGATATCATCTTGAAGACTGAAGTT 2181  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 |||||||  
 Db 2182 GAATAAGCCATCAGATGTCCGGAGCCGTATCAATGATGTTCCTGCTGATGACAC 2241

QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 |||||||  
 Db 2242 AGCCTAGATTTCTGGGGATACAGCAACACTTGGACCTCTCAACCAAGCCCTGTTTCC 2301  
 QY 741 IleTrpLeuIleValPheGlyValAlaMetGlyValIleValValGlyIleValIleLeu 760  
 |||||||  
 Db 2302 ATATGCTCATTTCTTTTGGAGTGTGCATGCGAGTGAATGCTTGGCATGTGCATCTCG 2361  
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro 780  
 |||||||  
 Db 2362 ATCTTCACTGGATCAGACATCGAGACAGAAATTAACCAAGAGTGAGAAATTCCT 2421  
 QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800  
 |||||||  
 Db 2422 TATGCTTCATCGATATATACCAAGGAAATATATCCAGAGATTCACAAACATGATGAT 2481  
 QY 801 ValGlnThrSerPhe 805  
 |||||||  
 Db 2482 GTTCAGACCTCCTT 2496  
 RESULT 4  
 AAD02758  
 ID AAD02758 standard; cDNA; 3396 BP.  
 AC AAD02758;  
 XX  
 DE 31-MAY-2001 (first entry)  
 XX  
 DE Human angiotensin converting enzyme-2 (ACE-2) cDNA.  
 XX  
 KW Human; angiotensin converting enzyme-2; ACE-2; peptidyl dipeptidase A;  
 KW screening; therapy; hypertension; congestive heart failure; CHF;  
 KW inflammation; pain; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FH Location/Qualifiers  
 FT 82..2499  
 FT /tag= a  
 FT /product= "Human angiotensin converting enzyme-2  
 FT (ACE-2)"  
 FT /EC\_number= "3.4.15.1"  
 FT /note= "This region is referred as SEQ.ID.NO.3 and is  
 FT specifically claimed in claim 26"  
 FT 82..135  
 FT sig\_peptide  
 FT /\*tag= b  
 FT 136..2496  
 FT mat\_peptide  
 FT /\*tag= c  
 FT /product= "Human mature angiotensin converting enzyme-2  
 FT (ACE-2)"  
 PN US6194556-B1.  
 XX  
 PD 27-FEB-2001.  
 PD  
 XX  
 PF 11-DEC-1997; 97US-0989299.  
 PF  
 XX  
 PR 11-DEC-1997; 97US-0989299.  
 PR  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 PA  
 PI Acton ST, Robison KE;  
 PI  
 XX  
 DR WPI: 2001-210604/21.  
 DR  
 DR P-PSDB; AAY72667.  
 DR  
 XX  
 PT Novel genes encoding angiotensin converting enzyme-2 useful as  
 PT antisense or antigenic agents for therapeutics, diagnostics and  
 PT screening assays -  
 PT  
 XX  
 PS Claim 1; Fig 1; 76pp; English.  
 PS  
 XX The present sequence is human angiotensin converting enzyme-2 (ACE-2)



CC CDNA. AE is also referred as peptidyl dipeptidase A. Nucleic acid  
CC sequence encoding ACE-2 is useful as antisense or antigenic agents for  
CC sequence specific modulation of gene expression or in the analysis of  
CC single base-pair mutations in the gene. Nucleic acid sequence encoding  
CC ACE-2 is useful in therapeutics, diagnostics and in screening assays.  
CC ACE-2 antagonist is used to treat hypertension or congestive heart  
CC failure (CHF). ACE agonist is used to reduce the inflammation and pain  
CC resulting from an insect sting or bite, which was accompanied by an  
CC injection of bradykinin. Anti-ACE-2 antibodies are used to monitor ACE-  
CC protein levels for determining the disease or condition associated with  
CC an aberrant protein level.

50 Sequence 3396 BP; 1034 A; 659 C; 772 G; 931 T; 0 other;

Alignment scores:	
Pred. No.:	0
Score:	4291.00
Percent Similarity:	100.00%
Best local Similarity:	100.00%
Query Match:	100.00%
DB:	22
Length:	3396
Matches:	805
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-978-385-2 (1-805) x AAD02758 (1-3396)

QY	1	MeSerSerSerSerTrpLeuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr	20
Dp	82	ATGTCAAACCTCTTCCCTGGCTCTCTTCAGCCCTTCTCTGTAACCTGCTGCACATCCACC	14
QY	21	IleGlnGlnAlaIleAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe	40
Dp	142	ATTGAGGAACAGGCCAAGACATTTTGTGACAAAGTTTAACCCGAGAGCCGAGACCTTTC	20
QY	41	TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGlnGluAsnValGln	60
Dp	202	TATCAAACTTCACCTGCTCTTCTGAAATTATACCCCAATATTACTGAAGAGAAATGTCCAA	26
QY	61	AsnMetAsnAsnAlaGlyAspLysTrpSerLysPheLeuLysGluGlnSerThrLeuAla	80
Dp	252	AACATGAAATTAATGCTGGGGGCAAAATGCTCTCTTTTAAAGGAACAGTCCACACTTGCC	32
QY	81	GlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu	100
Dp	322	CAAAATGATATCCACATCAAGAAATTCAGAAATCTCACAGTCAAGCTTCAGCTGCAGGCTCTT	38
QY	101	GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu	120
Dp	382	CAGCAAAATGGGCTCTTCAGTCTGCTCAGAAAGACAAGCAACAGCTTGAAACAATCTCA	44
QY	121	AsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnProGlnGlu	140
Dp	442	AATTCAAAGACACCACATCTACAGTACACGGAAGAAAGTTTGTAACCCAGATTAATCCACAAGAA	50
QY	141	CysLeuLeuLeuGlnProGlyLeuAsnGluIleLeuAlaAsnSerLeuAspTyrAsnGlu	160
Dp	502	TGCTATTACTTGAACCAAGGTTTAATGAATTAATGGCAAAAGCTTTAGACTCAACAAGAG	56
QY	161	ArgLeuTrpAlaTrpGlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyr	180
Dp	562	AGGCTCTGGGCTGGGAAGCTGAGATCTGAGGTGGCAAGCAGCGTACGCCATTATAT	62
QY	181	GlnGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly	200
Dp	622	GAAAGAGTATGGGCTTGAAAAAAGACATGGCAAGACCAATCATTTATGAGGACTATGGG	68
QY	201	AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly	220
Dp	682	GATTATTGGAGAGAGACATAGAAAGTAAATGGGGTAAATGGCTATGACTACAGCGCGGAGC	74
QY	221	GlnLeuIleGluAspValGluHisThrPheGlnGluTyrLysProLeuTyrGlnHisLeu	240
Dp	742	CAGTGTGATGAAGATGGGAACAATACCTTTGAAGAAAGTTAAACCATTTATATGAACATCTT	80
QY	241	HisAlaTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly	260

D	b	802	CATGCCTATGTGAGGGAAGTGTATGTAATGCCATTCCTCTCATATACGTCCATTTGGA	861
O	y	261	CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpHisLeuTyrSer	280
D	b	862	TGCCCTCCCTGCTCATTTTGCTGTGGATATGTGGGGTAGATTTTGGCAAAATCTGTACCT	921
O	y	281	LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln	300
D	b	922	TTGACAGTTCCTTTGGACAGCAAAACCAAAATAGATGTTACTGATGCATTTGGTGGACAG	981
O	y	301	AlaTrpAspAlaGlnArgIlePheLeysGlnAlaGlnLysPhePheValSerValGlyLeu	320
D	b	982	GCCTGGGATGCACAGAAATATTCAGAGAGCGCGAAGAAATCTTTGTATCTGTGGTCTT	1041
O	y	321	ProAsnMetTrpGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlnAsnValGln	340
D	b	1042	CCTATATGACTCAAGATTTCTGGAAATTTCCATCTTAACGAGACCCAGGAATGTTAG	1101
O	y	341	LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet	360
D	b	1102	AAAGAGCTGTCATCCACAGCTGTGGAGCTGGGGAAGGGGAGCTTCAGATTCCTATNG	1161
O	y	361	CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln	380
D	b	1162	TGCACAAAGGTGACAAATGGAGCGACTTCGACGGCTCATCATAGATGGGGCATATCCAG	1221
O	y	381	TyrAspMetAlaTyrAlaIleGlnProPheLeuArgAsnGlyValAsnGlnIlePhe	400
D	b	1222	TATGATATGGCATTTGCTGCACAACCTTTTCTGCTTAAGAAATGGAGCTATATAGAGATTC	1281
O	y	401	HisGlnAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer	420
D	b	1282	CATCAAGCTGTTGGGGAAATCATGATCACTTTCTGCGACCCACACTTAAGCATTTAAATCC	1341
O	y	421	IleGlyLeuLeuSerProAspPheGlnGlnAspAsnGluThrGlnIleAsnPheLeuLeu	440
D	b	1342	ATTGCTCTCTGTACCCGATTTTCAAGAGACAAATGAACAAATTAACCTCTGCTC	1401
O	y	441	LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg	460
D	b	1402	AAACAAGACACTCAGATTTGTTGGAGACTGTCCTCATTTACTTACATGTTAAGAAGTGGAG	1461
O	y	461	TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpLeuMet	480
D	b	1462	TGCGATGGCTTTAAAGGGGAAATTCCTCAAGACCATGGATTAATAAAGGTGGGAGATG	1521
O	y	481	LysArgGluIleValGlyValIleValGluProValProHisAspGluThrTyrCysAspPro	500
D	b	1522	AAAGGAGAGATAGTGGGGGTGGTGGAACTGTGCCCCCATGATGAANAATCATCTGTACCCC	1581
O	y	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu	520
D	b	1582	GCATCTCTGTGTCCATGTTTCTATATGATTTCTCATTCATTCGATTAACACAGGACCTT	1641
O	y	521	TyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis	540
D	b	1642	TACCAATTTCCAGTTTCAAGAGACATTTGTCAAGCAGCTTAACATGAAGGCCCTCTGCAC	1701
O	y	541	LysCysAspIleSerAsnSerTrpGlnAlaGlyGlnLysLeuPheAsnMetLeuArgLeu	560
D	b	1702	AAATGTGCATCTCAAACTCTACAGAAAGCTGGACAGAAACTCTTCAATATGTTAGGCTT	1761
O	y	561	GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsn	580
D	b	1762	GGAAAAATCAGAACCTTGAGACCTTAGCATTTGGAAAAATGTTGTGGAGCAAAAGCATGAT	1821
O	y	581	ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys	600
D	b	1822	GTAAGGCCACTGCTCAACTACTTTTGAGCCCTTAATTTACTGTGCTTAAGACACAGACAG	1881
O	y	601	AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal	620





QY	41	TyrIhbserSerIeuAlaSerTpaSnyrAsnThrAsnIlethnGluLysValGln	60
Db	202	TATCAAAAGTTCACCTGCTTCTTGGAATTATAACACCAATATTACGAGAGAAATGCCAA	261
QY	61	AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla	80
Db	262	AACATGAAATATAGTGTGGGACCAATGTGTCTGCTTTTAAAGAAACAGTCCACACTTCC	321
QY	81	GlnMetYrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu	100
Db	322	CAATATGATCCACTACCAAGAAATTCAGAAATCTCAGACGTCCAGCTTCAGGTGAGCTCTT	381
QY	101	GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu	120
Db	382	CAGCAAAATGGGTCTTCACGTGCTGCAGAAAGCAAGCAAAAGCGTTGACACAAATCTCA	441
QY	121	AsnThrMetSerThrIleYrSerThrGlnLysValCysAsnProAspAsnProGlnIu	140
Db	442	AATTCAAATGAGCACCACTACACGTACTGCGAAAGTTTGAACCCAGATATCCACAAATA	501
QY	141	CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspYrAsnGln	160
Db	502	TGCTATTACTTGACACAGGTTTGATGAAATATGCGAACAAAGTTTACACTCAATAG	561
QY	161	ArgLeuThrAlaTrpGlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuYr	180
Db	562	AGGCTCTGGGCTGGCAAACTGGAATCTGAGAGTGGCAACAGCTGAGCCATTATAT	621
QY	181	GluGluYrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspYrGly	200
Db	622	GAAAGATGTGGTGGTGAATAATGAGATGCGCAAGCAAGCAATATATAGAGCACTATGG	681
QY	201	AspYrTrpArgGlyAspYrGlyValAsnGlyValAsnGlyTyrAspYrSerArgGly	220
Db	682	GATTATTGGAGAGAGACGCTTGAAGATTAATGGGTGATGCTTATGACTACAGCCGCGC	741
QY	221	GlnLeuIleGluAspValGlnHisThrPheGlnGlnIleLysProLeuYrGlnHisLeu	240
Db	742	CAGTTGATTGAAGATGTGGAAACATACCTTTGAAGAAATTAACCATATATATGACATCTT	801
QY	241	HisAlaYrValArgAlaLysLeuMetAsnAlaTyrProSerYrIleSerProIleGly	260
Db	802	CATCCCTATGTAGAGGCGCAAGTGAATGAACTTACCTTCCATATACGTCCAAATTGGA	861
QY	261	CysLeuProAlaHisLeuLeuGlnAspMetTrpGlyArgPheTrpThrAsnLeuYrSer	280
Db	862	TGCTCTCCCTGCTCTTGGTGTGTGATATGTGGGTGATGATTTTGGACAAATCTGTACTCT	921
QY	281	LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln	300
Db	922	TTGACAGTCCCTTTGGACAGAAACCAACATATGATGTACTGATGCCAATGGTGGACAG	981
QY	301	AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu	320
Db	982	GCGTGGATGCGACAGAAATATTCAGAGAGGCGCAAGAACTCTTGTATCTGTGTGCTT	1041
QY	321	ProAsnMetThrGlnGlyPheTrpPGLuAsnSerMetLeuThrAspProGlyAsnValGln	340
Db	1042	CTTATATATGACTCAGAGATTTCTGGGAAATATCCATCTTAAACGACCCAGGAAATGTTCA	1101
QY	341	LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet	360
Db	1102	AAAGCAGTCTGCCATCCCAACGCTTGGAGACTGGGGAAGGCGACTTCAGGATCTTATG	1161
QY	361	CysThrLysValThrMetAspAspPheLeuThrAlaHisIleGlnMetGlyHisIleGln	380
Db	1162	TGCAACAAGGTGACATAGCAGCACTTCGACAGCTCATCATGAGATGGGCGCATATCCAG	1221
QY	381	TyrAspMetAlaYrAlaAlaGlnProPheLeuLeuArgAsnGlyValAsnGluGlyPhe	400
Db	1222	TATATATATGGCATATGCTGTGCACAACCTTTTCTGTAAAGAAATGAGCTATATAAAGATTC	1281

QY	401	HiGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuYsser	420
Db	1282	CATGAAGCTGTTGGGGAAATCATGTCACCTTTCGGACGCCACACCTTAAGATTAAATACC	1341
QY	421	IleGlyLeuLeuSerProAspPheGlnGluAspAspGluThrGluIleAsnPheLeuLeu	440
Db	1342	ATTGGTCTTCGTACCCGATTTTTCAAAGAAAGCAATGAAACAGAAATATAACTTCCTGCTC	1401
QY	441	LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg	460
Db	1402	AAACAGACACTCACGATTGTTGGAGCTGCGCAATTTACTTACATGTTAGAGAAAGTGAGAG	1461
QY	461	TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpProLue	480
Db	1462	TGGATGGCTTTAAAGGGAAATTCCTCAAGACCAAGCATGATATAAAGAGTGGAGATG	1521
QY	481	LysArgGluIleValGlyValIleGluProValProHisAspGluThrTyrCysAspPro	500
Db	1522	AAGCAGAGATAGTTGGGGTGGTGGAACTGTGCCCCATGATGAACATAGTGTACCCC	1581
QY	501	AlaSerLeuPheHisValIleSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu	520
Db	1582	GCATCTCGTTCATGTTTCATTAATGATTACTCATTCATTCGATATTACAGACAGACCTT	1641
QY	521	TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHis	540
Db	1642	TACCAATTCCTCAAGTTCAAAGAGCACTTTGCCAAGCAGCTAAACATGAAAGCCCTCGCAC	1701
QY	541	LysCysAspIleSerAsnSerThrGluAlaGlyIleLysLeuPheAsnMetLeuArgLeu	560
Db	1702	AAATGTGCATCTTAAACTCACGAAAGCTGGACACAAACTGTCTCAATGCTGAGGCTT	1761
QY	561	GlyLysSerGluProTyrPthrLeuAlaLeuGluAspValValGlyAlaLysAsnMetAsn	580
Db	1762	GGAAATACAGAACCTTGACCTTACGATTTGGAATAGTTGTGAGAGCAAGAACATGAT	1821
QY	581	ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys	600
Db	1822	GTAAGGCCACGCTCAACTACTTTGAGCCCTTATTACCTGAGCTGAAGACAGAACAG	1881
QY	601	AsnSerPheValGlyTyrPserThrAspTrpSerProTyrAlaAspGlnSerIleLysVal	620
Db	1882	AATCTTTTGGGATGGAGTACGACGATGGAGTCAATGACAGACCAAGCACTCAAACTG	1941
QY	621	ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluThrPasnAspAsnLue	640
Db	1942	AGGATAAACCTTAAATATAGCTCTTGGAGATAAAGCATATGATGACACACATTAATG	2001
QY	641	TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn	660
Db	2002	TACCTGTTCCGATCATCTGTTGCATATGCTATGAGCGAGACTTTTAAAGTAAATAAT	2061
QY	661	GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer	680
Db	2062	CAGATGATCTTTTGGGGAGAGAGATGGCGAGTGGCAATTTGAAACCAAGATCTCC	2121
QY	681	PheAsnPhePheValIThrAlaProLysAsnValSerAspIleIleProArgThrGluVal	700
Db	2122	TTTAATTTCTTTGCACCTGCACCTTAAATATGTCTCGATATCATTTCTTGAACTTAAGTT	2181
QY	701	GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn	720
Db	2182	GAATAAGGCCATCAGATGTCGCCGAGCGGTATCAATGATGCTTCGCTGTAATGACAC	2241
QY	721	SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer	740
Db	2242	AGCCTAGAGTTCGGGGAGTACAGCCAAACCTTGGACCTCTTACACAGGCCCTGTTTCC	2301
QY	741	IleTrpLeuIleValPheGlyValAlaMetGlyValIleValValGlyIleValIleLeu	760
Db	2302	ATATGGCTGATTTGTTTGGAGTTTGATGGAGTATATGTGGTGGCTTGTATCTCTG	2361
QY	761	IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro	780

Db	2362	ATCTCACTGGGATCAGATCGAAGAAAGAAAATAAAGCAAGAAGTGAGAAAATCT	2421
Oy	781	TyrAlaSerIleAspIleSerIysGlyGluAsnAsnProGlyPheGlnAsnThrAsp	800
Db	2422	TATCCCTCCATCGCATATATACCAAGAGAAAATAAATCCAGATCCAAAACATGATAT	2481
Oy	801	ValGlnThrSerPhe	805
Db	2482	GTCAGACCTCCTT	2496
RESULT 6			
AAD32586			
ID	AAD32586	standard; cDNA; 3396 BP.	
XX	AAD32586;		
XX			
XX	18-JUN-2002	(first entry)	
XX			
DE	Human ACE-2 full-length cDNA.		
XX			
KW	Human; angiotensin converting enzyme-2; ACE-2; therapy; hypertension;		
KW	peptidyl dipeptidase A; blood pressure; hypotension; atherosclerosis;		
KW	myocardial infarction; heart failure; arrhythmia; renal failure; gene;		
XX	inflammation; fertility; enzyme; X chromosome p21-22; ss.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	5'UTR	1..81	
FT	/*tag= a	82..2499	
FT	CDS	82..2499	
FT	/*tag= b	82..135	
FT	/product= "human ACE-2 protein"	82..135	
FT	/*tag= c	136..2496	
FT	mat_peptide	136..2496	
FT	/*tag= d	136..2301	
FT	/product= "human mature ACE-2 protein"	136..2301	
FT	/EC_number= "3.4.15.1"	136..2301	
FT	misc_feature	/*tag= e	
FT	/*note= "Encodes extracellular domain"	1201..1215	
FT	misc_feature	/*tag= f	
FT	/*note= "Encodes zinc binding domain (ZBD)"	2302..2376	
FT	misc_feature	/*tag= g	
FT	/*note= "Encodes transmembrane domain"	2377..2496	
FT	variation	/*tag= h	
FT	/*note= "Single chain polymorphism; This polymorphism results in an amino acid change from Asn to Asp"	2377..2496	
FT	misc_feature	/*tag= i	
FT	/*note= "Encodes cytoplasmic domain"	2500..3396	
FT	3'UTR	/*tag= j	
FT	variation	/*tag= k	
FT	/*note= "Single chain polymorphism"		
XX	WO200212471-A2.		
XX	14-FEB-2002.		
XX			
XX	09-AUG-2001; 2001WO-US25059.		
XX			
XX	09-AUG-2000; 2000US-0635501.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
DI	Acton S, Robison KE, Hsieh FY;		

XX	WPI: 2002-257481/30.
DR	P-FSDB; AAE20353.
XX	
PR	Isolated human polypeptide, known as angiotensin converting enzyme-2,
PT	useful for treating or preventing the development of an abnormal blood
PP	pressure or related diseases, e.g. hypertension, heart failure or
PN	myocardial infarction -
XX	
PS	Claim 1; Fig 1; 21bp; English.
XX	
CC	The invention relates to human angiotensin converting enzyme-2 (ACE-2)
CC	polypeptides and polynucleotides. ACE-2 is also known as peptidyl
CC	dipeptidase A (EC 3.4.15.1). Polypeptides of the invention are useful
CC	for treating or preventing the development of abnormal blood pressure
CC	and diseases or disorders associated with the protein in a subject. The
CC	diseases include hypertension, hypotension, congestive heart failure,
CC	chronic heart failure, acute heart failure, myocardial infarction,
CC	atherosclerosis, arrhythmia and renal failure. They are also useful
CC	for treating inflammatory conditions and diseases relating to fertility
CC	The present sequence is human full-length ACE-2 cDNA. The ACE-2 gene is
CC	located on the X chromosome at p21-22.
SQ	
	Sequence 3396 BP; 1034 A; 659 C; 772 G; 931 T; 0 other;
Alignment Scores:	
Pred. NO.:	0 Length: 3396
Score:	4291.00 Matches: 805
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	gaps: 0
US-09-978-385-2 (1-805) x AAD32586 (1-3396)	
QY	1 MetSerSerSerSerTTPLeuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
Db	82 ATGTCAACCTTTCTCGGCTCCTTCAGCCTTGTTGCTGAATCGTGCCTCAATCCACC 141
QY	21 TllegluglinalalalySthrPhelenuAsplysPheasnHlsJualagluAspleupe 40
Db	142 ATTGGAGAACAGGCCAAGACATTMTTGGACAAGTTTAACCCAGAAAGCCAGACTGTTC 201
QY	41 TYrGlnSerSerleuAlaSerTPasntyrAsnthrsnlierhrnglugluAsnvalGln 60
Db	202 TATCAAACTTTCACCTGCTCTTGGAAATTTAAACACCAATATTAAGTGAAGAATAAGTCCAA 261
QY	61 AsnmetAsenAnlaAglyAsplysTrpserAlapheluVsgJuglinSerThrLeuAla 80
Db	262 AACATGAATTAATGCTGGGGCAAAATGGTCTCCCTTTTAAAGGAGACAGTCCACTGGCC 321
QY	81 GlInMetYrProLeugIngluIleGlInasnThuThrVallysLeugInLeuGlnAlaleu 100
Db	322 CAATGATATCCACATACAAAGAAATTCAGAAATCTCACAGTCAAGCTTCAGCTCAGCTCTT 381
QY	101 GlInGlnasnGlySerSerValleuSerGluAsplysSerlysaArgleuAsnthrlleleu 120
Db	382 CAGCAAAATGGGCTCTCAGTCTCTCAGAAACCAAGACCAACGGTTGAACACAAATTCCTA 441
QY	121 AsnhrMetSerThrIleTySerThrhGlyslValCysasnProAspAsnProGlnGlu 140
Db	442 AATTCAAATGACACCATCTTACAGTACGTGAAAAGCTTTGTAAACCCAGATTAATCCCAAGAA 501
QY	141 CysteuleuLeuEngluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTYrAsnGlu 160
Db	502 TGCTTATTAATCTTGAACCAAGGTTTAATATGAATTAATGGCAACAGATTTTAGACTACAAAGAG 561
QY	161 ArgLeutrpAlatprGlnSerTerParSergluValalglyvsgInleuArqProLeutyrr 180
Db	562 AGGCTCTGGGCTTGGGAAGCTGAGANTCTGAGGTTCGGCAAGCAGCTGAGGCCATTATAT 621
QY	181 GlucgluYrValValleuLysAsnGluMetalaArglaAsnHlsYrgluAspTYrGly 200

Db 622 GAAGAGTATGTGCTTGAAAAATGAGATGCGAAGACCAATCATATGAGACTATGGG 681  
 QY 201 ASPTTTPATGAGLYASPTTGTGVALASNGLYVALASPGLYTTPASPTTYSERARGLY 220  
 Db 662 GATTATTTGGAGAGAGACTATGAAATGAGGGGTAGATGGCTATGACTTACAGCGCGGCG 741  
 QY 221 Glnleuileglnaspvalaighisthrpheglugluileysproleuetyrghisileu 240  
 Db 742 CAGTGATTTGAAGATGTGGAACATACCTTTAAGACAGATTAACCATTTATGAAACATCTT 801  
 QY 241 Hisalatyralargalalysleuemetasnalatyprosertrylieserproilegly 260  
 Db 802 CATGCCATGTGAGGGCAAGATTGATGATGCTTATCTCCATTCCTATACAGCAATTGGA 861  
 QY 261 CysleuproalaHisleuileuglyaspmetrprolyargphetrphtyrasnleutyser 280  
 Db 862 TGCCCTCCCTGCTCATTTGCTTGTGTGATATGTGGGGTAGATTTTGCAAAATCTGTACTCT 921  
 QY 281 Leuthrvalprophleglglulysproasnilleaspyalthrasspalametvalaspgl 300  
 Db 922 TTGACAGTTCCTTTGGAGCAAAACCAACATATGATGTACTGATGCAATGGTGGACCAG 981  
 QY 301 Alatrpspalaalnaragliepheylsglualeglulysphephevalservalglyleu 320  
 Db 982 GCCTGGGATGACAGAGATATTCAGAGAGGCGGAGAAAGTTCTTGTATCTGTGGCTT 1041  
 QY 321 Proasmetthrnglnghlypheprrgluansermetleuthraspprogllyasnvalgl 340  
 Db 1042 CCTATATGACTCAAGGATTTCTGGGAAATTCATGCTTACGCGGCCAGCAAAAGTTTCAG 1101  
 QY 341 LysalavalcysHisprothralatrpaspheuglylysglyasphearglieleuemet 360  
 Db 1102 AAACAGACTGCCATCCACAGCTTGGGACCTGGGGAAGGCGCACTTCAGAGATCTTATG 1161  
 QY 361 CysThrlysalthrmetaspaspheleuthralHisisglumetglyHisilegl 380  
 Db 1162 TGCAACAAGGTGACAAATGAGAGCACTTCTGCACAGCTCATATGAGTGGGCATATCCAG 1221  
 QY 381 TyraspmetalaTyralaialaglnprophleuenuargsnghlyalaasnghlyphe 400  
 Db 1222 TATATATGGCATATGCTGCAACACCTTTTCTGTAAAGAAATGAGGTAAATGAAAGATTC 1281  
 QY 401 Hisglualavalglyluilemetserleuseralalaatrrprolyshisleysser 420  
 Db 1282 CATGAACCTGTTGGGAAATCATGTCTTCTGCACGCCACACCTAAGCATTTAAATCC 1341  
 QY 421 Ileglyleuenuserproasppheglngluaspasnghlythrghlyleasnphleuenu 440  
 Db 1342 ATTGTCTTGTCTACCCGATTTTCAAGAAACATGATAAACAAGAAATTAACCTTCTGCTC 1401  
 QY 441 Lysglinalaleuthrilevaliglythrleuprophetrrymetleugluuystprarg 460  
 Db 1402 AAACAACACACTCAGATTTGTTGGAGCTCTGCCATTTACTTACATGTTAGGAAGTGGAGG 1461  
 QY 461 TrpMetvalpheylsglyluileprolyaspglatrrpmetlyslsyrtrprrglumet 480  
 Db 1462 TGGATGCTCTTAAAGGGGAAATCCCAAGACAGAGATGAAAGAAAGTGGTGGAGATG 1521  
 QY 481 LysarggluilevaliglyvalalguiprovalproHisaspeluthrtyrcysasppro 500  
 Db 1522 AAGGGAAGATAGTTGGGGGTGGGAACCTGTGCCCATATGATAAATCTCTGTGACCCC 1581  
 QY 501 AlaserleupheHisValaserasnasptryserpHeileargtryrThrargthrleu 520  
 Db 1582 GCATCTCTGTTCATGTTTCTAATGATTAATCATCTTATGATATTTACACAGAGACCTTT 1641  
 QY 521 TyrglnpheglnglnghlyalaaleucysglinalalaHisshisglnglyproleuHis 540  
 Db 1642 TACCAATTCAGATTCAAGAGCACTTGTCAAGACAGCTTAACATGATGAAGCCCTGTGAC 1701  
 QY 541 LyscysaspilIeserasnserthrnglualeglnglnlyseuHeasmetleuargleu 560  
 Db 1702 AATGTGACATCTCAAACTCTACAGAGAGCTGAGCAAGAAACTGTTCAATATGCTGAGGCTT 1761

QY 561 GlylyssercluprottrpHrleualaleuGluasnvalaiglyalalysasnmetasn 580  
 Db 1762 GGAATAATCAGAACCCCTGAGACCTTATGATGGAATAATGTTGTAGAGCAAAAGAACATGAT 1821  
 QY 581 ValArgProleuenuasntryrPheglupProleupheThrTripleuysaspghinslnys 600  
 Db 1822 GTAAAGCCACTGCTCAACTACTTTTGAGCCCTTATTTACTGGCTGAAAGACCAAGAACAG 1881  
 QY 601 AsnserphevalaiglytrpserthrAsptryserprotyralaspinsrileuysval 620  
 Db 1882 AATTCTTTGTGGAGATGAGGTACCGACTGGAGTCCATATGACAGCAACATCAAGATG 1941  
 QY 621 ArgIleleuileuysseralaleuGlyaspIysalaTyrgluTrpAsnaspasnghumet 640  
 Db 1942 AGGATAGGCTTAATAACACTCTTGGAGATTAAGCATATGAAATGCAACATCAATGAAATG 2001  
 QY 641 TyrlleupheargserSerValaIatyrAlametArgglntyrPheleuysValIysasn 660  
 Db 2002 TACCTGTTCCGATCATCTTGTGCATATGCTATGAGGAGCTACTTTTAAAGTTAAATAAT 2061  
 QY 661 GlumetIleleupheglngluaspvalArgvalalalaasnleuysproargIleler 680  
 Db 2062 CAGATGATTCCTTTTGGGAGAGATGTCGAGTGGCTAATTTGAAACCAAGATCTCC 2121  
 QY 681 PheasnPhephevalthrAlaprolysasnvalSeraspIleleProargThrghlyval 700  
 Db 2122 TTTAAATTTCTTGTCTACCTCACCTTAATAATGTTGTGATATCATCTTCAGAACTGAAGTT 2181  
 QY 701 GluIysAlaIleargmetSerArgserArgIleasnAspAlaPheargleuasnspasn 720  
 Db 2182 GAAAGAGCCATCAGAGATGTCGCCGAGCCGTATCATATGATGCTTCCGTGTATGATCAAC 2241  
 QY 721 SerleuGluPheleuglyIleGlnProthrleuglyProProasnghlyProProvalser 740  
 Db 2242 AGCCTGAGTTCTTGGGATACAGCCACACTTGGACCTCTTAACAGAGCCCTGTTCC 2301  
 QY 741 IleThrleuilevalPhegllyvalaIemetglyvalIlevalaIaiglyIlevalIleleu 760  
 Db 2302 ATATGGCTGATTTGTTTGGAGTGTGATGGAGAGATAGTGGTGGCATTTGTCATCTG 2361  
 QY 761 IlePheThrGlyIleargsparglyslslysnLysAlaArgSerGlygluasnPro 780  
 Db 2362 ATCTTACAGCGCATCGATCGAATCGAAGAAATTAAGCAAGAGTGGAGAAATCTT 2421  
 QY 781 TyralaSerIleaspIleSerlysglyluasnasnproglyPheglInasnThrAspasp 800  
 Db 2422 TATGCTTCATCATATTTAGCAAAAGAGAAATAATATCCAGATTCCAAAACACTGATGAT 2481  
 QY 801 ValGlnThrSerPhe 805  
 Db 2482 GTTCAGACCTCTT 2496  
 Db 2482 GTTCAGACCTCTT 2496  
 RESULT 7  
 AAS21279  
 ID AAS21279 standard; cDNA; 3732 BP.  
 XX  
 AC AAS21279;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cDNA sequence encoding for PRO1885 polypeptide.  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
 KW breast; prostate; cervical; tumour necrosis factor- $\alpha$ 1pha; TNF- $\alpha$ 1pha;  
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIa; gene therapy; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200140466-A2.  
 XX  
 PD 07-JUN-2001.



Db	1000	CGAATATGACGCCAAGATTTCTGGGAAAATTCATGCTAAACGAGACCCAGAAATGTTGAC	1055
Qy	341	LYSAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet	360
Db	1060	AAAGCAGCTGTGCCATGCCACAGCTGGGACCTGGGGAAGGGCACTTCAGATTCCTTATG	1119
Qy	361	CysThrLysValThrMetAspAspPheLeuThrAlaHisIleGluMetCylHisIleGln	380
Db	1120	TGCACAAAGGTGACATGACAGACTTCCTTAACAGCTCATCTGAGTATGGGATATCCAG	1179
Qy	381	TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe	400
Db	1180	TATGATATGGCATATGCTGTCACAAACCTTTCTCTGTAAGAAATGGACCTATATGAAAGTTC	1239
Qy	401	HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer	420
Db	1240	CATGAAAGCTGTGGGGAAATCATGCTACTTTCGCGACCCACACCTAAAGCATTTAAATACC	1299
Qy	421	IleGlyLeuLeuSerProAspPheGlnGluAspAsnGlyIthrGluIleAsnPheLeuLeu	440
Db	1300	ATTGGCTCTCTGTCACCGGATTTTCAAGAAGAAATGAACAAGAAATTAATCTCTGCTC	1359
Qy	441	LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg	460
Db	1360	AAACAGACAGCTACGATTTGGGACCTGCCATTTACTTACATGTTAGAGAGCTGAGAG	1419
Qy	461	TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpProGluMet	480
Db	1420	TGGATGGCTTTAAAGGGAAATTTCCCAAGAACCACTGATTAAGAAAGGGGGAGATG	1479
Qy	481	LysAspGluIleValGlyValValGluProValProHisAspGluThrTyrCysAspPro	500
Db	1480	AAGCGAGAGATAGTTGGGGGTGGGAACCTGGTCCCAATGATGAAGAACTACTGTGACCC	1539
Qy	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu	520
Db	1540	GCATCTCTGTCGATGTTTCTGTGATTACTCATCTTCATTCGATATTAACAAGAGCCCTT	1599
Qy	521	TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis	540
Db	1600	TACCAATTCAGTTTCAAGAGACCTTGTCAAGCAGCTAAACATGAAGAGGCCCTCTGCAC	1659
Qy	541	LysCysAspIleSerAsnSerThrGluLagIlyGlnLysLeuPhe-----	555
Db	1660	AAATGTGACATCTCAAACTCTACAGAACTGGACAGAAACCTTT-CTAAGAAATACCTCA	1718
Qy	555	-----	555
Db	1719	AAATGTTGAACCTCTCCTAGTATTCAGTATTACTCATTTCCATGCGCTAGTTGTATTGC	1778
Qy	555	-----	555
Db	1779	ATTTCCTTGTCTTAAAAAGAAATTTTATGCGCTTAAAAATGCTCTCATTTACAAACCAA	1838
Qy	555	-----	555
Db	1839	CATTTAATTTGTGTGACAGAGAACCTTAGACCATACAACAATTGGGTGGGCCACCTCTT	1898
Qy	555	-----	555
Db	1899	TTTCTCCATCATACTACAGCCCTCTCTTCTGTAAATTTGGAAGAAAGAGCGGTTTAA	1958
Qy	555	-----	555
Db	1959	GGTGGATATATCTGTTAATATGCACTTCTTTCTTATCTGCGCAACACCAATTTAGCCAA	2018
Qy	555	-----	555
Db	2019	GTCAAAGAGAGAAACCATAGATCATAGATGAATATATATATGTACATCTGGAACCCCTCAA	2078
Qy	556	-----AsnMetLeuArgLeuGluLysLysSerGluPro-----	565
Db	2079	AAGGCCCTGAACCCCTTTTTTTGTGTAGCAATATATGCTGAGAGCTTTGGAAATATGAAACCC	2138

[illegible]



XX	cachexial; male infertility; impotence; testicular cancer; lung tumour;
KW	hyperproliferative disorder; pulmonary system disorder;
KW	central nervous system disorder; bone disorder; Parkinson's disease;
KW	neurodegenerative disease; Alzheimer's disease; dementia; paraneoplasia;
KW	Huntington's disease; schizophrenia; mania; dementia; paraneoplasia;
KW	panic disorder; learning disability; amyotrophic lateral sclerosis;
KW	psychosis; autism; sleep disorder; immune system disorder;
KW	Hashimoto's thyroiditis; musculo-skeletal system disorders;
KW	multiple sclerosis; ischaemic brain injury; stroke; infectious disease;
KW	diabetes mellitus; immunological disorder; asthma; AIDS;
KW	acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;
KW	inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;
KW	neural system disorder; respiratory disorder; olfactory disorder;
KW	wound healing; chromosome X.
XX	
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	213..2348
FT	/tag= a
FT	/product= "NHP #1"
FT	/transl_except= (pos:867..869,aa:Xaa)
FT	/transl_except= (pos:930..932,aa:Xaa)
FT	/transl_except= (pos:1707..1709,aa:Xaa)
FT	/note= "Xaa= Any amino acid"
PN	
XX	W0200174896-A1.
PD	
XX	11-Oct-2001.
PF	
XX	02-APR-2001; 2001WO-US10542.
PR	
XX	03-APR-2000; 2000US-194118P.
PR	29-SEP-2000; 2000US-236384P.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	
PI	Moore PA, Ni J, Soppet DR, Coleman TA, Gentz RL, Endress GA;
PI	L1 Y, Dillon PJ;
XX	
XX	WPI: 2001-626394/72.
DR	P-PSDB: AAU09092.
XX	
PT	New human proteins, useful for diagnosing, treating, preventing and/or
PT	prognosing disorders related to the proteins, including cardiovascular
PT	disorders, autoimmune disorders and reproductive disorders -
XX	
PS	Claim 1; Page 291-292; 318pp; English.
XX	
CC	The invention relates to novel human proteins (NHP) and the
CC	nucleic acids that encode them and antibodies raised against them.
CC	The proteins, antibodies and nucleic acids are useful in the diagnosis,
CC	prognosis, prevention and/or treatment of diseases and/or disorders
CC	involving vasoconstriction, gastrointestinal disorders, cardiovascular
CC	disorders (e.g. hypertension, erectile dysfunction, high blood pressure,
CC	coronary heart disease and arteriosclerosis), anorexia, obesity, bulimia,
CC	cachexia, disorders of small intestine, disorders of reproductive system
CC	(e.g. male infertility and/or impotence), testicular cancer, lung tumours
CC	and other hyperproliferative disorders, disorders of pulmonary system,
CC	central nervous system disorders, bone disorders, neurodegenerative
CC	diseases and behavioural disorders (e.g. Alzheimer's disease, Parkinson's
CC	disease, Huntington's disease, schizophrenia, mania, dementia, paraneoplasia,
CC	panic disorder, learning disabilities, amyotrophic lateral sclerosis,
CC	psychoses, autism, sleep disorders), immune system disorders (e.g.
CC	Hashimoto's thyroiditis), renal and musculo-skeletal system disorders,
CC	central nervous system disorders (e.g. multiple sclerosis, ischaemic
CC	brain injury and/or stroke), infectious diseases, diabetes mellitus,
CC	immunological disorders (e.g. asthma, acquired immunodeficient syndrome
CC	(AIDS), leukaemia, rheumatoid arthritis, inflammatory bowel disease,
CC	sepsis, acne, psoriasis and lupus erythematosus), neural system
CC	disorders, respiratory disorders, olfactory disorders and wound
CC	healing. The present sequence encodes an NHP of the invention and
CC	is located on the X chromosome.



QY 302 PASPALAGlnArgIlePheLeuysGluAlaGluysPhePheValSerValGlyLeuProAs 322  
 Db 935 GGAGGACAGAGAAATATTCAAGAGAGGCCGAGAACTTCTTGATGCTGGTCTTCTAA 994  
 QY 332 nMethrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGlnLysAl 342  
 Db 995 TATAGCTCAAGGATTCGGGAAATTCATGCTAACGAGCAGGAAATGTTGAGAAAC 1054  
 QY 342 aValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMetCysTh 362  
 Db 1055 AGTGTGCATGCCACACCTTGGACCTGGGAAAGGCGACTTCAGGATCTTATGTGCAC 1114  
 QY 362 rLysValThrMetAspPheLeuThrAlaHisGluMetGlyHisIleGlnTrpAs 382  
 Db 1115 AAAGGTACAAATGAGACACTTCTGACAGCTCATGAGATGGGCAATATCCAGTATGA 1174  
 QY 382 pMetAlaTrpAlaGlnProPheLeuLeuAlaHisGluMetGlyHisIleGlnTrpAs 402  
 Db 1175 TATGGCATATGCTGCACAACTTCTGCTTAAGAAATGAGACCTAATGAGATTCATGA 1234  
 QY 402 uAlaValGlyLulIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSerIleG 422  
 Db 1235 AGCTGTGGGGAATCATGCTCACTTCTGCAGCCACACTTAAGCATTTAAATCCATTGG 1294  
 QY 422 yLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLysG 442  
 Db 1295 TCTTGTGTACCCGATTTTCAAGAAACATGAAACAGAAATTAATCTTCCGTCAACCA 1354  
 QY 442 nAlaLeuThrIleValGlyThrLeuProPheThrTyrrMetLeuGlyLysTrpArgTrpMe 462  
 Db 1355 AGACATCAGATGTTGGGACTCTGCCATTACTTACATGTTAGAGAGATGGAGTGAT 1414  
 QY 462 tValPheLysGlyLulIleProLysAspGlnTrpMetLysTrpTrpGluMetLysAr 482  
 Db 1415 GGTCTTTAAAGGGGAAATCCCAAGACAGATGATGAAAAAGTGTGGAGATGAAACG 1474  
 QY 482 gGluIleValGlyValGluProValProHisAspGluThrTyrrCysAspProAlaSe 502  
 Db 1475 AAGGATAGTGGGGTGGTGGGAACTGTCCCATGATATAAACATCTGTGACCCCGATC 1534  
 QY 502 rLeuPheHisValSerAsnAspTrpSerPheIleArgTyrrThrArgThrLeuArg 522  
 Db 1535 TCTGTTCATGTTCTAATGATTAATCTTATTCATTCATGATATACCAAGAGCCCTTACCA 1594  
 QY 522 nPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlnLysProLeuHisLysCy 542  
 Db 1595 ATTCCAGTTTCAAGAGACCTTTGTCAAGACAGCTAAACATGAAGGCCCTTCGCAAAATG 1654  
 QY 542 sAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGly 562  
 Db 1655 TCACATCTCAAACTCTACAGAACTGGACAGAACTGTCAATATGCTGAGGNTTGGAAA 1714  
 QY 562 sSerLeuProTrpThrLeuAlaLeuGluAsnValValGlyValAlaLysAsnMetAsnValAr 582  
 Db 1715 ATCAAGAACCCGAGACCTAGCATTTGAAAAAGTTGTGAGGAGCAAGAACATGATGTAA 1774  
 QY 582 gProLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSe 602  
 Db 1775 GCCACTGCTCAACTTGTAGCCCTTATTTACTGCTGAAAGACAGCAAGAAATTC 1834  
 QY rPheValGlyTrpSerThrAspTrpSerProTyrrAlaAsnGlnSerIleLysValArgI 622  
 Db 1835 TTTTGTGGATGGATGCCAGCTGAGATCCATATGACAGACCAAGCATCAAAAGTGAGAT 1894  
 QY eSerLeuLysSerAlaLeuGlyAspLysAlaTyrgLutTrpAsnAspAsnGluMetTyrrLe 642  
 Db 1895 AAGCCTAAATCAGCTCTTGAGATTAAGCATATGATGAGAACGACAAATGATGACCT 1954  
 QY uPheArgSerSerValAlaTyrrAlaMetArgGlnTyrrPheLeuLysValLysAsnGluMe 662  
 Db 1955 GTTCGATCATCTTGTCCATATGCTATGAGGACGACTCTTTTAAAGATAAAATCATGAT 2014  
 QY tIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProArgIleSerPheAs 682

Db 2015 GATTCCTTTTGGGAGAGAGATGCGAGCGCTAATTTGAACCAAGATCTCCCTTAA 2074  
 QY 682 nPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGly 702  
 Db 2075 TTTTGTGTACCTGACCTCAAAAATGTCTGTATCATCTTCAGAACTGAGTTGAAA 2134  
 QY 702 sAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLe 722  
 Db 2135 GCCCATCAGATGCTCCCGGAGCCGTATCATATGATGCTTCCGTGGAAGACAGACCT 2194  
 QY 722 uGluPheLeuGlyTrpIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTr 742  
 Db 2195 AGAGTTTCTGGGATACAGCCAACTTGAGACTCTTAACGAGCCCGCTTTCATATG 2254  
 QY 742 pLeuIleValPheGlyValValMetLysValIleValValGlyIleValIleLeuIlePh 762  
 Db 2255 GCTGATTTGTTTGGAGTTGTGATGGAGTGAATGATGGCATTCATCTGATCTT 2314  
 QY 762 eThrGlyIleArgAspArgLysLys 770  
 Db 2315 CACTGGATCAGAGATCGAAGAG 2339

## RESULT 9

AA514890 standard; cDNA: 2911 BP.

AA514890;

20-DEC-2001 (first entry)

Human cDNA encoding novel human protein NHP #11.

KW Human; novel human protein; NHP; ss; antidiabetic; antirheumatic;  
 KW antiarthritic; cytostatic; antiarteriosclerotic; vulnerrary;  
 KW neuroprotective; nootropic; antiparkinsonian;  
 KW anti-human immunodeficiency virus; antilastimatic; vasotropic; cardiant;  
 KW antitumic; anorectic; antileptility; neuroleptic; anticonvulsant;  
 KW antinflammatory; antibacterial; antiparasitic; thyromimetic;  
 KW immunomodulator; antidiabetic; dermatological; vasoconstriction;  
 KW gastrointestinal disorder; cardiovascular disorder; hypertension;  
 KW coronary heart disease; arteriosclerosis; anorexia; obesity; bulimia;  
 KW cachexia; male infertility; impotence; testicular cancer; lung tumour;  
 KW hyperproliferative disorder; pulmonary system disorder;  
 KW central nervous system disorder; bone disorder;  
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;  
 KW Huntington's disease; schizophrenia; mania; dementia; paranoia;  
 KW panic disorder; learning disability; amyotrophic lateral sclerosis;  
 KW psychosis; autism; sleep disorder; immune system disorder;  
 KW Hashimoto's thyroiditis; musculo-skeletal system disorders;  
 KW multiple sclerosis; ischaemic brain injury; stroke; infectious disease;  
 KW diabetes mellitus; immunological disorder; asthma; AIDS;  
 KW acquired immunodeficient syndrome; leukaemia; rheumatoid arthritis;  
 KW inflammatory bowel disease; sepsis; acne; psoriasis; lupus erythematosus;  
 KW neural system disorder; respiratory disorder; olfactory disorder;  
 KW wound healing.

Homo sapiens.

Key Location/Qualifiers

Key 213..998

CDS /tag- a

/product- "NHP #11"

WO200174896-A1.

11-OCT-2001.

02-APR-2001; 2001MO-US10542.

03-APR-2000; 2000US-194118P.

29-SEP-2000; 2000US-236384P.



KW	Alzheimer's disease; hormone; cytokine.
OS	Homo sapiens.
XX	
PN	JPL1318472-A.
XX	
PD	24-NOV-1999.
XX	
PF	22-JAN-1999; 99JP-0014949.
XX	
PR	13-MAY-1998; 98GB-0010373.
PA	18-AUG-1998; 98GB-0018009.
XX	
XX	(SMIK ) SMITHKLINE BEECHAM PLC.
DR	WPI: 2000-109268/10.
XX	
PT	MPROT15 polypeptide and MPROT15 polynucleotides - useful for the treatment of hypertension, myocardial diseases, apoplexy, heart diseases, nervous denaturation, Alzheimer's disease etc.
XX	
PS	Claim 18; Page 15; 22pp; Japanese.
CC	
CC	This is coding sequence #2 of human MPROT15. The MPROT15 polynucleotide and polypeptide sequences can be used for the treatment of hypertension, myocardial diseases, apoplexy, heart diseases, nervous denaturation, Alzheimer's disease and diseases related to the processing of peptide hormones and cytokines.
XX	
SQ	Sequence 2262 BP; 693 A; 450 C; 523 G; 596 T; 0 other;
Alignment Scores:	
Pred. No.:	0
Percent Similarity:	3740-50
Best Local Similarity:	90.57%
Query Match:	89.43%
DB:	21
US-09-978-385-2 (1-805) x AAZ59466 (1-2262)	
QY	11 LeuValAlaValThrAlaAlaGlnSerThrIleGlnGlnAlaAlaLysThrPheLeuAsp
Db	64 CTGTGGCTGTAACGTCGTCTACGTCACCCATTGAGGACAGCCAGCATTTTGGAC
QY	31 LysPheAsnHisGluAlaGluAspLeuPheTyrGlnSerSerLeuAlaSerTTPAsnTyr
Db	124 AAGTTTAACCGACGAAGCCGAAGACCTGTTCTATCAAAAGTTCACCTTGTGGAAATTAT
QY	51 AsnThrAsnIleThrGlnGluAsnValGlnAsnMetAsnAsnAlaGlnLysAspLysTTPSer
Db	184 AACACCATATTACTGAAGAGAATGTCCAAACATGATAATAGTGGGCAAAATGGTCT
QY	71 AlaPheLeuLysGlnGlnSerThrLeuAlaGlnMetTyrProLeuGlnGlnIleGlnAsn
Db	244 GCCTTTAAAGGAACAGTCCACACTGCCCCAAATGTATCCACTACAGAAATTCAGAAAT
QY	91 LeuThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlnSerSerValLeuSerGln
Db	304 CTCACAGTCAAGCTTCAGCTCAGAGGCTCTTCAGCAAAATGGTCTTCAGTGTCAAGAA
QY	111 AspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrngly
Db	364 GACAAAGACCAACCGTTGAACACACATTCCTAAATACATGAGCACCATCTCACACTGGA
QY	131 LysValCysAsnProAspAsnProGlnGlnCysLeuLeuLeuGlnProGlnLysLeuAsnGln
Db	424 AAATTTGTATACCGACAGTAAATCCACAGAGATGCTTATTTACTTACACCAAGTTTGAATGAA
QY	151 IleMetAlaAsnSerLeuAspTyrAsnGluArgLeuThrPalaIlePheLysSerTTPAspSer
Db	484 ATATATGCCAAACAGTTTAGACTACAAATGAGAAGGCTGTGGCTTGGGAAAGCTGAGATATCT

QY 171 GlnValGlyLysGlnLeuArgProLeuTyrGlnGluTyrValValLeuLysAsnGlnMet 190  
 |||||  
 Db 544 GAGTCGGCAGACGAGTGGAGCCATATATGAAAGATGATGGCTTGAAGAAAAGAGATG 603  
 QY 191 AlaArgAlaAsnHisTyrGlnAspTyrGlyAspTyrTrpArgGlyAspTyrGlnValAsn 210  
 |||||  
 Db 604 GCAAGAGCAAAATCATTTATGAGCACTATGGGATTTATGGAGAGAGCACTATGAAGTAAAT 663  
 QY 211 GlyValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGlnAspValGlnHisThrPhe 230  
 |||||  
 Db 664 GGGGTAGATGGCTATGATACAGCCGGCCAGTTGATGAAAGATGGAAACATACCTTT 723  
 QY 231 GlnGluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsn 250  
 |||||  
 Db 724 GAAGGATTTAAACCATTTATGAAACATCTTCATGCTATGAGAGGCAAACTTATAAAT 783  
 QY 251 AlaTyrProSerTyrIleSerProIleGlyLysLeuProAlaHisLeuLeuGlyAspMet 270  
 |||||  
 Db 784 GCGTATCTCTCTATATCATAGTCAATGGATGGCTCCCTGCTCATTTGGCTGTGATATG 843  
 QY 271 TrpGlyArgPheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 290  
 |||||  
 Db 844 TGGGTAGATTTTGGCAAAATCTGTACTCTTGGACAGTTCCCTTGGACAGAAACCAAC 903  
 QY 291 IleAspValIleThrAspAlaMetValAspGlnAlaTrpAspAlaGlnArgIlePheLysGln 310  
 |||||  
 Db 904 ATGATGTCTTACTGATGCAATGTGTGACCAAGCTGTGGATGCAAGCAATTTTCAGAGAG 963  
 QY 311 AlaGluLysPhePheValSerValIleLeuProAsnMetThrGlnGlyPheTrpGluAsn 330  
 |||||  
 Db 964 GCCGAGAGTCTTGTGTCTGTGTCTCTCTATATGATGATCAAGATCTCGGGAAT 1023  
 QY 331 SerMetLeuThrAspProGlyAsnValGlnLysAlaValLysHisProThrAlaTrpAsp 350  
 |||||  
 Db 1024 TCCATGCTTACGAGCCAGCAAAATGTTCAGAAAGCACTGTCCATCCACAGCTTGGGAC 1083  
 QY 351 LeuGlyLysGlyAspPheArgIleLeuMetCysThrLysValIleThrMetAspPheLeu 370  
 |||||  
 Db 1084 CTGGGGAAGGGGAGCTTCCAGATCTCTTATGTGCACAAAGTGACAAATGAGACACTCTCG 1143  
 QY 371 ThrAlaHisHisGlnMetGlyHisIleGlnTyrAspMetAlaTyrAlaAlaGlnProPhe 390  
 |||||  
 Db 1144 ACAGCTCATCATGAGATGGGGCATTCAGATGATGATGATGATGATGATGATGATGATGATG 1203  
 QY 391 LeuLeuArgAsnGlyAlaAsnGlnGlyPheHisGlnAlaValIleGlyIleMetSerLeu 410  
 |||||  
 Db 1204 CTGCTTACAGAAATGAGCTTATGAAAGATTCCTCATGAACTGTGGGGAATCATGCACTT 1263  
 QY 411 SerAlaAlaIleThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGlnGlu 430  
 |||||  
 Db 1264 TCTGAGCCACACCTTAAAGATTTAAATCCATTTGCTCTTGTCCACCGATTTTCAGAA 1323  
 QY 431 AspAsnGluThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIleValIleGlyThrLeu 450  
 |||||  
 Db 1324 GACCAATGAAACAGAAATAACTCTGCTCAACAGACATCAGATGTTGGGACTGTG 1383  
 QY 451 ProPheThrTyrMetLeuGlnLysTrpArgTrpMetValIlePheLysGlyIleProLys 470  
 |||||  
 Db 1384 CCATTTACTTACATGTTAGAGAGAGTGGAGTGGATGCTTTTAAAGGGGAATTTCCCAA 1443  
 QY 471 AspGlnTrpMetLysLysTrpTrpGluMetLysArgGluIleValIleValIleGluPro 490  
 |||||  
 Db 1444 GACAGGTGATGAAAGAGTGGGAGATGAA----- 1476  
 QY 491 ValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAspTyr 510  
 |||||  
 Db 1476 ----- 1476  
 QY 511 SerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnGlnAlaLeuLys 530  
 |||||  
 Db 1477 -----TATTACCAAGAGACCTTTTACCAATTCACATTTCCAGAACACATTTGT 1524  
 QY 531 GlnAlaAlaLysHisGlnGlyProLeuHisLysCysAspIleSerAsnSerThrGluAla 550

Db 1525 CAAAGCCTTAAACATGAAAGCCCTTGGCAAAATGTGACATGTCAAACTTACAAAGCT 1584  
 |||||  
 QY 551 GlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProTrpThrLeuAlaLeu 570  
 |||||  
 Db 1585 GGACAGAAACTGTTCAAATGTGCTGAGAGCTTGGAAATATGAAACCTGGACCTTACCATTTG 1644  
 QY 571 GlnAspValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGluPro 590  
 |||||  
 Db 1645 GAAATGTGTAGAGAGCAAGAACATGAAATGTAAGCCACTGCTCACTACTTTTAAGCC 1704  
 QY 591 LeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValIleTyrSerThrAspTrp 610  
 |||||  
 Db 1705 TTATTTACCTGGCTGAAGACCAAGAACAGAAATTTCTTTGTGGATGAGTACCACTGG 1764  
 QY 611 SerProTyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAlaLeuLysPhe 630  
 |||||  
 Db 1765 AGTCCATGG----- 1773  
 QY 631 LysAlaTyrGluThrPasnAspAsnGlnMetTyrLeuPheArgSerSerValAlaTyrAla 650  
 |||||  
 Db 1774 -----GAACTTTCATCTCTGATTTGTGCTCTGGCA 1809  
 QY 651 MetArgIleTyrPheLeuLysValLysAsnGlnMetIleLeuPheGlyGlnLysAspVal 670  
 |||||  
 Db 1810 CAAGTGAAGATGTTGTGTT-----TTGTTCTCTACAGGAGGAGGAGATGTG 1854  
 QY 671 ArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheValIleThrAlaProLysAsn 690  
 |||||  
 Db 1855 CGAGTGGCTTAATTTGAAACCAAGAAATCTCTTAATTTTGTGTCTGCTGCTACCTTAAAT 1914  
 QY 691 ValSerAspIleIleProArgThrGluValGluLysAlaIleArgMetSerArgSerArg 710  
 |||||  
 Db 1915 GTGTGATATCATTTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1974  
 QY 711 IleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGlnPheLeuGlyIleGlnProThr 730  
 |||||  
 Db 1975 ATCAATGATCTCTTCCGTGCAATGACCAAGCTTGAAGTCTGGGATGATGACCAAGCA 2034  
 QY 731 LeuGlyProProAsnGlnProProValSerIleTrpLeuIleValPheGlyAlaValMet 750  
 |||||  
 Db 2035 CTGGAGCCTCTTAAACAGCCCTGTTTCCATTTGCTGATTTTGGAGTGTGATG 2094  
 QY 751 GlyValIleValValGlyIleValIleLeuIlePheThrGlyLleArgAspArgLysLys 770  
 |||||  
 Db 2095 GAGATGATGATGTTGGCATTTGCATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 2154  
 QY 771 LysAsnLysAlaArgSerGlyLysAsnProTyrAlaSerIleAspIleSerLysGlyLys 790  
 |||||  
 Db 2155 AAAATATTAACCAAGAACTGGAAGAAATCCTTATGCTCCATGCAATTTTACCAAGAGAGA 2214  
 QY 791 AsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPhe 805  
 |||||  
 Db 2215 AATATATCCAGATTTCCAAACACATGATGTTCCAGACCTCTTT 2259  
 |||||  
 RESULT 11  
 AAC84368  
 ID AAC84368 standard; cDNA; 2638 BP.  
 XX  
 AAC84368;  
 AC  
 XX  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 XX  
 DE  
 XX  
 XX  
 Mouse zace2-5 protein encoding cDNA.  
 zace2: metalloenzyme; angiotensin-converting enzyme; ACE; fertility;  
 znc metalloproteinase; blood pressure; znc protease; hypertension;  
 ventricular systolic dysfunction; renal impairment; heart failure;  
 scleroderma renal crisis; atherosclerosis; antiinflammatory; mouse;  
 antiatheritic; bradykinin inactivator; ss.  
 XX  
 XX  
 OS Mus sp.

Key Location/Qualifiers  
 FT CDS 106..2523  
 FT /\*tag= a  
 FT /product= "Zace2-5"  
 FT /note= "the coding fragment is specifically claimed for"  
 XX  
 PN W0200070032-AL.  
 XX  
 PD 23-NOV-2000.  
 XX  
 PE 03-MAY-2000; 2000WO-US11932.  
 XX  
 PR 13-MAY-1999; 99US-0311482.  
 PR 27-AUG-1999; 99US-0384706.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Pladdington CS, Petrie CR, Shoemaker KE, Bishop PD;  
 XX  
 DR WPI; 2001-025018/03.  
 DR P-PSDB: ABA48097.  
 XX  
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory  
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases  
 PT associated with inflammation such as arthritis and enterocolitis -  
 XX  
 PS Claim 10; Page 104-109; 125pp; English.  
 XX  
 CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-  
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood  
 CC pressure regulation and fertility. Zace2 can be expressed by standard  
 CC recombinant methodology. Zace2 polypeptides are useful for treating an  
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),  
 CC diseases associated with inflammation like arthritis and enterocolitis),  
 CC as targets for identifying modulators of zinc protease activity, for  
 CC screening or identifying new angiotensin-converting enzyme (ACE)  
 CC inhibitors, and as a basis for rational drug design for inhibitory  
 CC molecules. The nucleic acids can be used to detect the expression of a  
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and  
 CC for detecting and localizing Zace2 gene expression in tissue samples,  
 CC to determine whether a subject's chromosomes contain a mutation in the  
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.  
 CC Inhibitors of ACE are used for treating hypertension of various  
 CC conditions, including left ventricular systolic dysfunction, progressive  
 CC renal impairment, scleroderma renal crisis, congestive heart failure due  
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be  
 CC used to treat infertility while Zace2 antagonists are used for inducing  
 CC infertility. The present sequence represents a cDNA encoding the mouse  
 CC Zace2-5 protein.  
 XX  
 SQ Sequence 2638 BP; 802 A; 556 C; 611 G; 669 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0 Length: 2638  
 Score: 3579.00 Matches: 661  
 Percent Similarity: 89.578 Conservative: 60  
 Best Local Similarity: 82.118 Mismatches: 84  
 Query Match: 83.418 Indels: 0  
 DB: 22 Gaps: 0  
 US-09-978-385-2 (1-805) x AAC64368 (1-2638)  
 QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20  
 DB 106 ATGTCCAGCTCCCTCCGCTCTCTCAGCTTGTGTACTGCTAGTCCGCTC 165  
 QY 21 ILGLUGLUGLAlaLysThrPheLeuAspLysPheAsnHisGlnLysGlnAspLeuPhe 40  
 DB 166 ACCGAGGAAATATCCCAAGACATTTTAAACAACTTAAATCAAGAACTCAAGCCGTCT 225  
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGlnLysAsnValGln 60  
 DB 226 TATCAAGAGTTCACTGCTCTTGTGATTAATTAATCACTTAAGTCAAGAAATGCCCA 285

QY 61 AsnMetAsnAlaLysLysTrpSerAlaPheLeuLysGlnLysSerThrLeuAla 80  
 DB 286 AAGATGAGTAGAGCTGCTACACCCAAATGCTGCTTTTAAAGAAACAGTCAAGACTGCC 345  
 QY 81 GlnMetTyrProLeuGlnLysLeuGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100  
 DB 346 CAAAGTTCTCACTACAAAGAAATCCGACTCCGATCATCAACGCTCACTACAGGCTCT 405  
 QY 101 GlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrIleLeu 120  
 DB 406 CACCAAGAGTGGCTCTTACACACTCTCAGCAGCAAGAAACAGTGAACCAATCTCTG 465  
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140  
 DB 466 AACCATATGAGCACCATTATACAGTCTGTAAGAAAGTTTGCACCCCAAGACCCCAAGAA 525  
 QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGlu 160  
 DB 526 TGTCTTATTACTTGAGCCAGGATTTGATGAATTAATGCGCACAGCAACAGTCACTACTCT 585  
 QY 161 ArgLeuTrpAlaTrpGlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180  
 DB 586 AGGCTCTGGGCAATGGAGGCTGAGGCTGAGGTTGGCAACGACTGAGGCCCTTGTAT 645  
 QY 181 GlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspTyrGly 200  
 DB 646 GAGAGATATGCTGTACTGTAAGAAACAGATGGCAAGCAACAACTTAATACCATATGGG 705  
 QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220  
 DB 706 GATTAATGAGAGGGGACATATACACAGAGGAGCAGATGGTACAACTATTAACGCTAAC 765  
 QY 221 GlnLeuIleGlnAspValGlnHisThrPheGlnGlnIleLysProLeuTyrGlnHisLeu 240  
 DB 766 CAGTTGATGAGAGATGTAAGAACGTACTCCGACAGATCAACCATGTATGAGACTCTT 825  
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleLys 260  
 DB 826 CATGGCTATGATGAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 885  
 QY 261 CysLeuProAlaHisLeuLeuGlnAspMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280  
 DB 886 TCCCTCCCTGCCCATTTGCTGTGATATGTTGGTATGTTGGACAAATCTGACCT 945  
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 DB 946 TTGACTGTTCCTTTCACAGAAACCAATAGATGTTACTGATGATGATGATGATGATGAT 1005  
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGlnLysGlnLysPhePheValSerValGlyLeu 320  
 DB 1006 GCGTGGATGACGAAGAAGATATTTCAAGAGCAGAGAAATCTTGTCTTGTGGGCTT 1065  
 QY 321 ProAsnMetThrGlnGlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsnValGln 340  
 DB 1066 CCTCATATGACCAAGAGATTTGCGGCAAACTATGCTGACTGACCCAGCAGATGGCCGG 1125  
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 DB 1126 ABAAGTTGTGTGACCCCAAGCTTGGATCTGGGACAGGAGACTTCAGATCAAGATG 1185  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlnHisIleGln 380  
 DB 1186 TGTACAAAGGCTACAAATGAGACACTTCTTGAAGGCCATTCAGAGATGGGACATCCAA 1245  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyLysPhe 400  
 DB 1246 TATGACATGAGCATATGACAGGCAACCTTCTCTTAAGAAACGAGCAATGAAGAGGCTTC 1305  
 QY 401 HisGluAlaValGlyGlnIleMetSerLeuSerAlaAlaTrpProLysHisLeuLysSer 420  
 DB 1306 CATCAAGCTGTGGAGAAATCATGCTACTTCTGAGCTACCCCAAGCATCTGAATATCC 1365

QY 421 ILeGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440  
 |||||  
 DB 1366 ATGGCTCTTGCCATCCGATTTTCAGAGATAGCCGAACAGATGAACCTTCTCTG 1425  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrlMetLeuGluLysTrpArg 460  
 |||||  
 DB 1426 AACAGGCAATGACAAATGTTGGAACACTACCGTTTACTTACATGTTAGAGAACTGAGG 1485  
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpGluMet 480  
 |||||  
 DB 1486 TGGATGGCTTTCGGGGGGAATTCGCAAGAGCAGTGAGGAAAGTGGTGGAGATG 1545  
 QY 481 LysArgGluIleValGlyValIleValGluProValProHisAspGluThrTyrlCysAspPro 500  
 |||||  
 DB 1546 AAGCGGAGATCGTTGGTGTGTGGAGCCCTGCTCTCATGATGAACATACCTGACCCCT 1605  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrlSerPheIleArgTyrlThrArgThrLeu 520  
 |||||  
 DB 1606 GCATCTCTGTTCCATGTTTCAATGATTAATCATTCATTCGATTTACACAGAGCATCT 1665  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGluGluAlaAlaLysHisGluGlyProLeuHis 540  
 |||||  
 DB 1666 TACCAATTCAGATTCAAGAGAGCTCTTGTCAAGCAGCTAGTATATGTTCTCTGAC 1725  
 QY 541 LysCysAspIleSerAsnSerThrGluIleValGluGluLysLeuPheAsnMetLeuArgLeu 560  
 |||||  
 DB 1726 AATATGCAATCTCAATATTCACATGAGCTGGGCAAGAGTTGCTCAGATCTGAGCTT 1785  
 QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsn 580  
 |||||  
 DB 1786 GGAATTCAGAGCCCTGGAGCCAAAGCTTGGAAATGTGTAGGAGCAAGAAATATGAT 1845  
 QY 581 ValArgProLeuLeuAsnTyrlPheGluProLeuPheThrTrpPheLysAspGluAsnLys 600  
 |||||  
 DB 1846 GTAACACCACTGCTCAATTAATCTCCACACGTTGTGACTGGCTGGAAGACAGACAGCA 1905  
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrlAlaAspGlnSerIleLysVal 620  
 |||||  
 DB 1906 AATTTCTTTTGTTGGGTGGAGACCTGATGAGCCCATTTGCCGACCAAGCATTAAGTG 1965  
 QY 621 ArgIleSerLeuLysSerAlaLeuGluLysPylsAlaTyrlGluTrpAsnAspAsnGluMet 640  
 |||||  
 DB 1966 AGGATTAAGCCCTAAATCAAGCTCTGAGCTATGATGATGATGAGCAACAGCAAGATG 2025  
 QY 641 TyrlLeuPheArgSerSerValAlaTyrlAlaMetArgGlnTyrlPheLeuLysValLysAsn 660  
 |||||  
 DB 2026 TTCCCTGTTCCGATCAATGTTTCATATGCAAGAAAGATTTTCAATATCAAAATAC 2085  
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 |||||  
 DB 2086 CAGACAGTCCCTTCTTAGAGAAATGATGAGAGTGCAGCATTTCAACCAAGAGTCC 2145  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700  
 |||||  
 DB 2146 TTCTACTCTTCTTGTACCTCACCCCAAAATGTCGTGATGATCTCTAGAGAGGAAATG 2205  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 |||||  
 DB 2206 GAAGATGCCATCAGATGCTGCGGCGCATCAATGATGCTTGGCCCTGGAATGATAC 2265  
 QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 |||||  
 DB 2266 ACCCTGGAGATTCTGGGATTCACCCAACTAGCCACCTTACACACCTCTCTGTCAC 2325  
 QY 741 IleTrpLeuIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeu 760  
 |||||  
 DB 2326 AATAGCGCATATATTTGCTGTGTGATGACACTGTAGTGCTGATCATCATCATCTG 2385  
 QY 761 IlePheThrGlyIleArgAspArgLysLysAsnLysAlaArgSerGlyGluAsnPro 780  
 |||||  
 DB 2386 ATTTGTCACCTGGATCAAGAGTGAAGAAAGAAATGAACCAAGAAAGAGAGAACCT 2445  
 QY 781 TyrlAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800

DB 2446 TATGACTCGATGGACATTTGAAAAGAGAAAGCAATGACGATTCAAAACAGTATGAT 2505  
 QY 801 ValGlnThrSerPhe 805  
 |||||  
 DB 2506 GCTCAGACTTCCTTT 2520  
 RESULT 12  
 AAC84370  
 ID AAC84370 standard; cDNA: 2638 BP.  
 XX  
 AC AAC84370;  
 XX  
 DT 19-MAR-2001 (first entry)  
 XX  
 DE Mouse Zace2-10 protein encoding cDNA.  
 KW Zace2; metalloenzyme; angiotensin-converting enzyme; ACE; fertility;  
 KW zinc metalloproteinase; blood pressure; zinc protease; hypertension;  
 KW ventricular systolic dysfunction; renal impairment; heart failure;  
 KW scleroderma renal crisis; atherosclerosis; antiinflammatory; mouse;  
 KW antithrptic; bradykinin inactivator; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 106..2523  
 FT /tag="a  
 FT /product="Zace2-10"  
 PN WC0200070032-A1.  
 XX  
 PD 23-NOV-2000.  
 XX  
 PF 03-MAY-2000; 2000MO-US11932.  
 XX  
 PR 13-MAY-1999; 99US-0311482.  
 PR 27-AUG-1999; 99US-0384706.  
 XX  
 PA (ZYMO ) ZYMOGENETICS INC.  
 XX  
 PI Piddington CS, Petrie CR, Shoemaker KE, Bishop PD;  
 XX  
 DR WPI: 2001-025018/03.  
 DR P-PSDB: AAB48098.  
 PT Angiotensin-converting enzyme, Zace2, useful for treating inflammatory  
 PT bowel disease, e.g. Crohn's disease and ulcerative colitis, or diseases  
 PT associated with inflammation such as arthritis and enterocolitis -  
 XX  
 PS Disclosure: Page 113-118; 125pp; English.  
 CC The invention relates to the metalloenzyme Zace2. Zace2, an angiotensin-  
 CC converting enzyme is a zinc metalloproteinase that plays roles in blood  
 CC pressure regulation and fertility. Zace2 can be expressed by standard  
 CC recombinant methodology. Zace2 polypeptides are useful for treating an  
 CC inflammatory bowel disease (e.g. Crohn's disease and ulcerative colitis),  
 CC diseases associated with inflammation like arthritis and enterocolitis,  
 CC as targets for identifying modulators of zinc protease activity, for  
 CC screening or identifying new angiotensin-converting enzyme (ACE)  
 CC inhibitors, and as a basis for rational drug design for inhibitory  
 CC molecules. The nucleic acids can be used to detect the expression of a  
 CC Zace2 gene in a biological sample, as probes for in vivo diagnosis and  
 CC for detecting and localizing Zace2 gene expression in tissue samples.  
 CC to determine whether a subject's chromosomes contain a mutation in the  
 CC Zace2 gene, and to detect aberrations associated with the Zace2 locus.  
 CC Inhibitors of ACE are used for treating hypertension of various  
 CC conditions, including left ventricular systolic dysfunction, progressive  
 CC renal impairment, scleroderma renal crisis, congestive heart failure due  
 CC to dysfunction, and treatment of atherosclerosis. Zace2 agonists may be  
 CC used to treat infertility while Zace2 antagonists are used for inducing  
 CC infertility. The present sequence represents a cDNA encoding the mouse  
 CC Zace2-10 protein.



XX Sequence 2638 BP; 797 A; 557 C; 615 G; 669 T; 0 other:

## Alignment Scores:

Pred. No.:	0	Length:	2638
Score:	3561.00	Matches:	659
Percent Similarity:	89.32%	Conservative:	60
Best Local Similarity:	81.86%	Mismatches:	86
Query Match:	82.99%	Indels:	0
DB:	22	Gaps:	0

US-09-978-385-2 (1-805) x AAC84370 (1-2638)

```

QY 1 MetSerSerSerSerTrpLeuLeuLeuSerLeuValAlaValThrAlaIleGlnSerThr 20
Db 106 ATGTCCAGGTCCTCTGGGCTCTTCACACCTTGCTGCTTACTCTGCTCAGTCCCTC 165
QY 21 IleGluGluGlnAlaIleAlaThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
Db 166 ACCGAGGAAATGCCAAGACATTTTAAACAACTTTAAACAGAGGCTGAAACACCTGTCT 225
QY 41 TyrGlnSerSerLeuAlaSerTrpAsnThrAsnIleThrGluGluAsnValGln 60
Db 226 TATCAAAAGTTCACCTGCTCTTGGATTTATTAATACTAATTAAGTGAAGAAATGCCCA 285
QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 286 AAGATAGTAGAGGCTGCACCAAAATGCTGCTCTTTATGAAAGAACAGCTAAAGCTGCC 345
QY 81 GlnMetTrpProLeuGlnIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100
Db 346 CAAAGTTCCTCACTACAAATAATCCAGACTCCGATCATCAACGGCTACACAGCCCTT 405
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
Db 406 CAGCAAAAGGGGCTTCACACCTCTCAGACAGCAAGAACAAACATTTGACACCAATTCTG 465
QY 121 AsnThrMetSerThrIleIleYserThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 466 AACACCATGAGCACCATTTTACAGTACTGTAAGTTTGGCAACCCCAAGAACCCCAAGAA 525
QY 141 CysLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTrpLysGln 160
Db 526 TGCCTTATTAATCTGACCCAGGATTTGATGAATAATAGCCGACACAGCAACATCAACTCT 585
QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTrp 180
Db 586 AGGCTCTGGGCGATGGAGGGCTGGAGGCTGAGTTGGCAAGACAGCTGAGGCCCTGTAT 645
QY 181 GlnGluTrpValValLeuLysAsnGlnIleMetAlaArgAlaAsnHisIleGlnAspTrpGly 200
Db 646 GAAAGCTATGTGGTCTGTAATAAAGATGGCAAGAACAAACAAATTAATACGACTATGGG 705
QY 201 AspTrpTrpArgGlyAspTrpGluValAsnGlyValAspGlyTrpAspTrpSerArgGly 220
Db 706 CATTAATGGAGGGGAGCTATGAAGCAGAGGAGAGAGATGGCTACAACTATTAACCGTATAC 765
QY 221 GlnLeuIleGluAspValGlnHisThrPheGluGlnIleLysProLeuTrpGlnHisLeu 240
Db 766 CAGTGTGATTGAAGATGTGAACGTACCTCCAGAGATCAACCACTATGATGAGCAGCTCT 825
QY 241 HisAlaTrpValArgAlaLysLeuMetAsnAlaTrpProSerTrpIleSerProIleGly 260
Db 826 CATGCTATGTGAGAGAGAAATGTGATGATACCTACCTCTCAATCAACGCCCACTGGA 885
QY 261 CysLeuProAlaHisIleLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTrpSer 280
Db 886 TGCCCTCCCTGCCCATTTGCTGTGGATATGTGGGGTATTTTGACAAATCTGTACCTT 945
QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 946 TTGACTGTTCCTTTGGACAGAAACCAAAACATATAGATGTACTGATGAATGATGAATCAG 1005

```

```

QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320
Db 1006 GGCTGGATGTCAGAAAGATATTTTCAGAGGACAGAAATCTTTGTTCTGTGGCCTT 1065
QY 321 ProAsnMetThrGlnGlyPheThrPglLysAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 1066 CCTCATATGACTCCAAAGATTTGGGGCAAACTTATGTGACTGAGCGACGCAAGATGGCCGG 1125
QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1126 AAAATTGTCTGCCACCCACACAGCTTGGGATCTGGGACACGGAGACTCAGAAATCAAGATG 1185
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisSHISGluMetGlyHisIleGln 380
Db 1186 TGTACAAAGGTTCACAAATGACAACTTTTGACACCCCATCCAGAGATGGGACACATCCAA 1245
QY 381 TyrAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
Db 1246 TATGACATGCGCATATGCCAGGCAACCTTCTGCTAAGAAACGAGCCATGAAGGGTTTC 1305
QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1306 CATGAAGCTGTGGAGAAATCATGTCACTTCTGACAGTACCCCAAGCATCTGAATCC 1365
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440
Db 1366 ATGTGCTTCTGCCATCCGATTTTCAAGAAATGCGAAACAGAGAACTCTTAATCC 1425
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTrpMetLeuGluLysTrpArg 460
Db 1426 AAACAGCATGTGAATGTTGGAACACTACCGCTTACTTAACATGATGAAGAGTGGAGG 1485
QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnIlePheLysLysTrpTrpGluMet 480
Db 1486 TGTATGGCTCTTGGGGGTGAAATTCCTCAAGAGCAGTGGATGAAATGCTGGTGGAGATG 1545
QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTrpCysAspPro 500
Db 1546 AAGGGGAGACGTTGGTGTGTGTGAGACCTGCTGCTGTATGAACATCTGTACCTT 1605
QY 501 AlaSerLeuPheHisValSerAsnAspTrpSerPheIleArgTrpTrpThrArgThrLeu 520
Db 1606 GCACCTCTGTCCATGTTTCAATGATTAATCACTTCAATGATTAACACAGACCATAT 1665
QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGluGlnAlaAlaLysHisGluGlyProLeuHis 540
Db 1666 TACCAATTCAGATTTCAAGAGGCTTTGTCAAGCAGCTAAGATATGATGTTCTGTGCAC 1725
QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560
Db 1726 AAATGTGACATTCCAATTCACATGGAAGCTGGGCAAGATGTCTCAAGATGCTGAGCTTT 1785
QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValAlaGlyAlaLysAsnMetAsn 580
Db 1786 GGAATTCAGAGCCCTGGACCGAAGCTTGGAAATGTGTAGAGACCAAGATATGAGAT 1845
QY 581 ValArgProLeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
Db 1846 GTAACACCATGCTCATTTCTTCCAAGCCTTTTGTGACTGCTGAAAGGCGAAGCAACA 1905
QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
Db 1906 AATTTCTTTTGGGGTGGAGACACTGAATGAGCCCATATGCCAGCAACCAAGATTAAGTG 1965
QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640
Db 1966 AGGATTAAGCTAAATACGCTTTGGAGCTAATGCAATTAATGAGCCCAACCAAGAAATG 2025
QY 641 TyrLeuPheArgSerSerValAlaTrpAlaMetArgGlnTyrPheLeuLysValLysAsn 660
Db 2026 TTCCCTGTCCGATCATCTGTGATATGCGATGAGAAATATATTTCTTAATATCAAAAC 2085
QY 661 GlnMetIleLeuPheGlyGluGlnAspValArgValAlaAsnLeuLysProArgIleSer 680

```





```

Db 601 GAVTAAATGGMNGNGNAGATAYAGARGTNAAYGGNGTGAAGTATGATATAYASNMNGCN 660
QY 222 HlGleuIIeGluAspValIuHlstrPhheglugluIIeLysProLeuTygluHlSteu 240
Db 661 CARYTNATGAGAGATGTNARCAACNTTYTGARGARATHAARCCNTYNTAYGARCAAYTN 720
QY 241 HlSalatYrValaArgAlaLysLeuMeLAsnaLatYrProSerTyrlIeserProIlegly 260
Db 721 CAYGCTAAGTMMGNCNNAARYTNMTGAAGCNTATCCWMTATATATHSMCCNATHGGN 780
QY 261 CysLeuProAlaHlSteuLeuLysPmetTyrgIyAArgPhetTyrlAsnLeuTyrlSer 280
Db 781 TGYTYNCCGNCAYTNTNGNCAATGTGGGNGMGTYYTGACNAAYTYNTAYMSN 840
QY 281 LeuThrValProPhheglYglInLysProAsnIleAspValLThrAspAlaMeValaSpGln 300
Db 841 YTNACNGTNCNTTYGGNARARARCCNNAAYTHGAYGNACGATGCAATGGTNGAYCAR 900
QY 301 AlATrPaspAlaGlnArgIlePhelYsgluAlaGluLysPhePheValSerValGlyLeu 320
Db 901 GCNTGGGAYGCCNARMGNATHTTYAARGARCCNARARATTYTYGTWMSGTNGCNTN 960
QY 321 ProAsnMetThrGlnGlyPheThrGluAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 961 CCNAAATGACNCARAGNTTYTGGAARAAYSNATGYTNACNGAYCCGNGNAAAYGTNCAR 1020
QY 341 LysAlaValCysHlSProThrAlaTrPaspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1021 AARGCNTGTGCAACNACNGCNTGGAYTNTGNGNARAGNGAYTMYGNAHYTHNATYG 1080
QY 361 CysTrIlySValThMeLAspAspPheLeuThrAlaHlSgluMetGlyHlSleGln 380
Db 1081 TGYACNARGTNACNATGGAATYTYTNTNACNCCNCAAYGARATGGGCAVATHCAR 1140
QY 381 TyrAspMetAlaYrAlaAlaGlnProPheLeuLysArgAsnGlyAlaAsnGlyLysPhe 400
Db 1141 TAYGATAGGCTNAYGCGNCARCCNTTYTNTNMGNAAYGGNGCAAGGAGNTTY 1200
QY 401 HlSgluAlaValGlyLulIeMetSerLeuSerAlaAlaThrProLysHlSleLysSer 420
Db 1201 CAYARGCNTGTNGNARATHTATGMSYTNMNSGNCNCCNNAARCAAYTNAARMSN 1260
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPhelLeu 440
Db 1261 ATHGCTNTNTMNSGCCGATTYTCARGARCAAYARARCCNARATHTAAYTYTYNTN 1320
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrlMetLeuGluLysTrpArg 460
Db 1321 AARCARCNTNMACNATHGNTGNAACNTTTCNATATATGTYNGARAAATGTGAGN 1380
QY 461 TrpMetValPheLysGlyLulIeProLysAspGlnTrpMetLysLysTrpGluMet 480
Db 1381 TGGATGGTNTTYAARGNGARATHCNARAGCAFTGATGAARARATGTTGGAGATG 1440
QY 481 LysArgGluIleValGlyValGluProValProHlSAspGluThrTyrlCysAspPro 500
Db 1441 AARMGNARARATHTGNGNNGTNGNARCCNCCNCAAYAGARACNTATYTYGAYCCN 1500
QY 501 AlaSerLeuPheHlSAsnAspTyrlSerPheIleArgTyrlTrpTrpArgThrLeu 520
Db 1501 GCNMSNTNTNTTYCAVTNMSNAAYGATATWMTTYTATHTGNTATYRNYACNMGACNTN 1560
QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHlSgluGlyProLeuHlS 540
Db 1561 TATCARTTYCARTTYCARGARCNNTNTGTCARCGNCCNARCAAYGARCGNCCNTNCAV 1620
QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyLysLeuPheAsnMetLeuArgLeu 560
Db 1621 AARTGCAVTATHTMSNAAYWMSNACNARCCNCGNCCARARATNTTYAATGTYTNMGNTN 1680
QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsn 580

```

---

```

Db 1681 GGNAAHMSNGARCCNTGACNTYNGCNTNGARAAATGTNGTNGNCCNAAARAAATGAAY 1740
QY 581 ValArgProLeuLeuAsnTyrlPhegluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
Db 1741 GTMNGCCNTYNTNAAATATATTYTGARCCNTNTTYACNCGTGTNNAARCAAYCARAAAR 1800
QY 601 AsnSerPheValGlyTyrlSerThrAspTrpSerProTyrlAlaAspGlnSerIleVal 620
Db 1801 AAYMSNTTYGTNGTNGTMSNACNGATYGMSNCCNATAYCCNCAAYCARMSNATHARGTN 1860
QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysValaTyrlGluThrPaspAsnAspGluMet 640
Db 1861 MGNATHMSNTNNAARMSNGCNTYTGNGNATAAARCCNTATYARKGGAAYGATYAGARATG 1920
QY 641 TyrLeuPheArgSerSerValaTyrlAlaMetArgGlnTyrlPheLeuLysValLysAsn 660
Db 1921 TAYTYNTTYMGMSNMSNGNTNGCNTAYGCNATGMCNARATYTYTNAARGTNARAAV 1980
QY 661 GlnMetIleLeuPheglYglInGluAspValaArgValaAlaAsnLeuLysProArgIleSer 680
Db 1981 CARATGATHTNTTYTGNGARGARAGATYTMGNGTNCNNAAYTNAARCCNMGNATHMSN 2040
QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
Db 2041 TYTAAATYTYTYGTNACNCCNCAARAAATCTMWSNCAATHTATHTCCNMGACNARGTN 2100
QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2101 GARAARCCNATHGNAAGMSNMSNMGNAATHAAGAYGCNTTYMGNTYNAAYGAYAAV 2160
QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyLysProAsnGlnProProValSer 740
Db 2161 WSNATNARATYTYTGNGNATHCARCCNACNTYTGNGCCNCCNNAAYCARCCNCCNGTMSN 2220
QY 741 IleTrpLeuIleValPheglYvalaMetGlyValaIleValaGlyIleValaIleLeu 760
Db 2221 ATHTGCTNATHTGNTTYGNGTNGTNTATGGGNGTNAHTGNTGNTGNTGNTGNTHTN 2280
QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
Db 2281 ATHTTYACNNGNATHTMGNGATYMGNAARAAARAAARAAAGCMWMSNCGNARAAAYCCN 2340
QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAspProGlyPheGlnAsnThrAspAsp 800
Db 2341 TAYGCMWSNATHTGATYTHMSNARAGNGARAAAYAAATCCNCGNTTYCARAAATCCNATGAY 2400
QY 801 ValGlnThrSerPhe 805
Db 2401 GTNCCARACNMSNTTY 2415

```

RESULT 14  
 ID ABK30270/C  
 ID ABK30270 standard; cDNA; 2350 BP.  
 ABK30270;  
 23-APR-2002 (first entry)  
 Human G-protein-coupled protease #40.  
 Human; ss; gene; G-protein-coupled protease; gene therapy;  
 transgenic; protease mediated disorder; proliferative disorder;  
 kw differentiative disorder; developmental disorder;  
 haematopoietic disorder;  
 Homo sapiens.  
 OS  
 PN US631427-B1.  
 XX  
 PD 18-DEC-2001.  
 XX  
 PF 26-MAR-1999; 99US-0280116.  
 XX

PR 26-MAR-1999; 9905-0280116.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 XX Robison KE;  
 PT  
 XX WPI; 2002-129545/17.  
 PS  
 XX  
 CC The invention relates to an isolated human protease nucleic acid molecule  
 CC comprising a nucleotide sequence of 546 base pairs, one of 268 fully  
 CC defined in the specification. Also disclosed are production of an  
 CC isolated polypeptide encoded by the nucleic acid, comprising introducing  
 CC the nucleic acid into a host cell and culturing under conditions to  
 CC express the protein from the nucleic acid, use of an antibody to  
 CC detect the encoded protein in a sample and to modulate its in vivo  
 CC activity, identifying agents that bind to the protein and identification  
 CC of a polynucleotide agent that modulates the expression of the nucleic  
 CC acid or its complement (i.e. gene therapy). The nucleic acid can be used  
 CC to identify an agent that modulates the expression or activity of the  
 CC nucleic acid, and can be used to isolate the protein. The nucleic acid  
 CC can be used in diagnostic assays for determining nucleic acid expression  
 CC as well as activity in the context of a biological sample (e.g., blood,  
 CC serum, cells, tissue) to determine whether an individual has a disease or  
 CC disorder, or is at risk of developing a disease or disorder, associated  
 CC with aberrant expression or activity of the nucleic acid. The nucleic  
 CC acid can be used to detect mutations in protease genes and gene  
 CC expression products such as mRNA. The nucleic acid can be used as  
 CC a hybridisation probe to detect naturally-occurring genetic mutations in  
 CC a protease gene. The nucleic acid can be used in drug screening methods  
 CC to identify agonists and antagonists that can be used to diagnose and  
 CC treat such protease mediated disorders e.g., proliferative,  
 CC differentiative, developmental or haematopoietic disorders. The nucleic  
 CC acid can be used as probes, primers, in biological assays, to determine  
 CC patterns of gene expression, to design ribozymes and to construct  
 CC transgenic animals. The present sequence represents one of the 268  
 CC disclosed human G-protein-coupled protease cDNA sequences.  
 XX  
 XX  
 SQ Sequence 2350 BP; 603 A; 548 C; 489 G; 710 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 7.28e-311 Length: 2350  
 Score: 3291.00 Matches: 726  
 Percent Similarity: 92.62% Conservative: 2  
 Best Local Similarity: 92.37% Mismatches: 12  
 Query Match: 76.70% Indels: 46  
 DB: 24 Gaps: 2  
 US-09-978-385-2 (1-805) x ABR30270 (1-2350)  
 QY 1 MetSerSerSerSerTirPleuLeuSerLeuValAlaVal-ThrAlaIaIaInSerTh 20  
 DB 2305 AAGTCAAGCTCTTCTGCTCTCTCTCAAGCTTGTCTGTAGATGCTGCTCAATCCAC 2246  
 QY 20 rllgluglgluAlaIa 39  
 DB 2245 CATTGAGAACAGCCCAACATTTTGGACAAAGTTTAAACCAACCAAGCCGAAGACTG 2186  
 QY 40 PheTYrGLaSerSerLeuAlaSerTirPasnTYrAsnThr-AsnIleThrGLuIaAsnVa 59  
 DB 2185 TTCTATCAAGTTCACCTCTCTTGTGAATTTAAACCCATATTACTGAAGAATATGT 2126  
 QY 59 IGLaAsnMetAsnAlaIa 78  
 DB 2125 CCAAAACATGAATTAATGCTGGGACAAATATGCTGCTTTTAAAGGAACAGTCCACA 2066  
 QY 79 LeuAlaIaIaMetTYrProIeuInIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 98  
 |||||||

DB 2065 CTGGCCCAATATGATATCACTACCAAGAAATTCAGAAATTCACAGTCAACCTTCATCTGCAG 2006  
 QY 99 AlaleuGLInGLaInGLySerSerValIeuserGLuAspIySerIyArgIeuaSnThr 118  
 DB 2005 GCTCTTCAACAAAATGGGCTTCAGTCTCTCAGAAAGCAAGAGCAAAAGGTTGAACACA 1946  
 QY 119 IleIeuaSnThrMetSerThrIle-TyrSerThrGLy-LysValCysAsnProAspAsn 138  
 DB 1945 ATTCTTAATACAAATAGACCACTCATACAGTCTGAAAAAGTTGTATACCAATATATC 1886  
 QY 138 roGLInGLuCysLeuLeuLeuGLuProGLyLeuAsnGLuIleMetAlaAsnSerLeuAsp 158  
 DB 1885 CACAAGAAATGCTTATCTTACCTGAACACAGCTTGAATGAATATGCAAAACAGTTTACCT 1826  
 QY 158 YrAsnGLuArgIeuaIa 178  
 DB 1825 ACATATAGAGGCTCTGGGCTTGGGAAAGTGAAGATCTGAGGTCGGCAAGACTGAGGC 1766  
 QY 178 roLeuTYrGLuGLuTYrValIaIeuaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 198  
 DB 1765 CATTTATTAAGAGATATGTGCTTGAATAAATGAGATGGCAAGCAAAATCATTTAGAG 1706  
 QY 198 sPTyGLyAspTYrTYrPArgIyAspTYrGLu-ValAsnGLyVal-AspGLyTYrAspTY 217  
 DB 1705 ACTATGGGATTAATGGAGAGAGACTATGAAGAAAGGGGTAAAGTCTATGACTA 1646  
 QY 217 r-SerArgIyGLuIleuIleGLuAspValGLuHisThr-PhaGLuGLuIleIyProIeu 236  
 DB 1645 CATGGCCGGCCAGTGTATGAAGATGGCAATACCTTTGAAGAGATTAACCATTA 1586  
 QY 237 TYrGLuHisLeuHisAlaTYrValArgAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 256  
 DB 1585 TATGAACATCTTCATCTCATGTGAGGCCAAAGTTGATAAAGCCATCTTCATATATC 1526  
 QY 257 SerProIleGLyCysLeuProAlaHisIleu-LeuGLyAspMetTYrGLy-ArgPheTYr- 275  
 DB 1525 AGTCCATTTGATGATCCCTCCCTGCTCATTTCCGTTGATATGTGGGCTTAATTTGGG 1466  
 QY 276 ThrAsnLeuTYrSerLeuThrVal-ProPheGLyIn-LysProAsnIle-AspValThr 294  
 DB 1465 ACAATCTGTACTTGTGACAGTGTCCCTTTGGACGGAACCAACCAATACATGTACT 1406  
 QY 295 AspAla-MetValAspGLuAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 313  
 DB 1405 GATGCCAATGTGTGACCAAGCCCTGGATGACCAAGAAATATTCAGAGGCCCGCAAG 1346  
 QY 314 PhePheValSerValGLyLeuProAsnMetThrGLInGLyPheTYr-GLuAsnSerMetIle 333  
 DB 1345 TTCTTTGTATCTGTGTGCTCTCTTAATATGACATCAAGATTTGGGAAATTTCCATGCT 1286  
 QY 333 uThr-AspProGLyAsnValGLuIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 352  
 DB 1285 AACGGACCCAGAAATGTTCAAGAAAGCACTGCCATCCACAGCTTGGGAGGGGG 1226  
 QY 353 LysGLyAsp-PhaArgIle-LeuMetCysThrIyAspMetAlaTYrAlaIaIaIaIaIaIa 372  
 DB 1225 AAGGGCCACTTCAAGATCTCTTAATATGACCAAGGTGACATGACAGCTTCTGACAG 1166  
 QY 372 IahIshIaIaMetGLyHisIa 392  
 DB 1165 CTGATCATGAGATGGGCAATATCCAGATGATGATGATGATGATGATGATGATGATGATG 1106  
 QY 392 euaTYrAsnGLyAlaAsnGLuIyPheHisGLuAlaIaIaIaIaIaIaIaIaIaIaIaIa 412  
 DB 1105 TAAAGAAATGAGCTTAAAGAAAGATTCATGACCTGTGGGAAATCATGTCACTTTCTG 1046  
 QY 412 IaaIa 430  
 DB 1045 CAGGCAACCTTAAGCAATTTAAATCCATTTGTTTGTACCCGAATTTTCAAGAA 986  
 QY 431 AspAsnGLuThrGLuIleAsnPheLeuLeuIyGLIn-AlaLeuThrIleValGLyThr-L 450  
 DB 985 GACAATGAACAGAAATAACTTCTCTCAAAACCAAGCACTACGATGTTGTTGGACTCC 926  
 |||||||



QY 40 PheTyrGlnSerSer-LeuAlaSer-TrypAsnTyrAsnThrAsnIlethrgInGluAsnV 59  
 |||||  
 Db 159 TTTCTATCAAAAGTTCAAGTTCCTTGGAAATATATACACCAATATATCTAGAGAAATG 218  
 QY 59 aIGln-AsnMetAsn---AsnAlaGlyAspLys-----TTPSerAlaPheLeuLysGlu 75  
 |||||  
 Db 219 TCCAAACAAATCATTAAGTTGCTGCGAGCAAAATGTGTCTAGCCCTTTTACAGGAA 278  
 QY 76 GlnSerThrLeuAlaGln-MetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLe 95  
 |||||  
 Db 279 CAGTCCACACTTGGCCCAAGATGTATCCATCAAA-----GCCAACTTCACGACATCTCC 332  
 QY 95 uGlnLeuGlnAlaLeuGlnGlnAsnGly-----SerSerVal 107  
 |||||  
 Db 333 ACATGTCAACGCTTCAGCTGTCCAGCGCTTCTTCAAGCCATTAACCTGTAGCTCTGAG 392  
 QY 107 IleuSer-----GluAspLysSerLys-ArgLeuAsnThrIleLeuAsnThrMet---- 123  
 |||||  
 Db 393 TTGGTCATCAGACGACAGACAGCAAAACCGGTGACACAAATTTCTATATACAAATGG 452  
 QY 124 -----SerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGluCysL 142  
 |||||  
 Db 453 AGCCCAACCAATCCTAACAGTAACTGAAACAGTCTGTAACCCAGATATACACAGAAATGCT 512  
 QY 142 eulLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGluArgL 162  
 |||||  
 Db 513 TATTACTTGAACAGAGTTTGAATGAATATGCAAAAGTTTATAGACTACATGAGAGGC 572  
 QY 162 eutThrAlaThrGlnSerThrPargSerGluValGlyLysGlnLeuArgProLeuTyrGlnG 182  
 |||||  
 Db 573 TCTGGGCTTGGGAAAGCTGAGATCTGAGTGGCAAGAGCTGAGGCCATTATATATAGAG 632  
 QY 182 IuTyrValValLeuLysAsnGlnMetAlaArgAlaAsnIleTyrGluAsp-Tyr-GlyAs 201  
 |||||  
 Db 633 AGTATGTGGTCTTGAATAATGAGATGGCAAGCAAAATCATATATAGAGCTTATGGGGA 692  
 QY 201 pTyrTrpArgGlyAspTyrGluValAsnGlyValAsp---GlyTyrAspTyrSerArgG 220  
 |||||  
 Db 693 TTAATGAGAGAGACTATGAGTAAATGGGTAAATAGTATGATATGATTAAGCCGCGG 752  
 QY 220 yGlnLeuIleGlnAspValGlnHisThr-PheGlnGlnIleLysProLeuTyr-GlnHis 239  
 |||||  
 Db 753 CCAGTTGATTTGAAGATGTGGAACATACCTGTTGAAGAGATTAACCATTTGATAGAGACT 812  
 QY 240 LeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIle 259  
 |||||  
 Db 813 CTTCAGCCCTATGTAGGAGCCCAAGTTGATGATGCCCTATCTCTATATCTGCATAT 872  
 QY 260 GlyCysLeuProAlaHisLeuLeuGlnLysAspMetTrp-GlyArgPheTrpThrAsnLeu 279  
 |||||  
 Db 873 GGAATGCCCTCCCTGCTCATTTGGTGGTATGTGCGGGTAGATTTTGGACAAATCTGTA 932  
 QY 279 rSerLeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAs 299  
 |||||  
 Db 933 CTCTTTGACAGTCCCTTTGGACAGAAACCAACATAGATGTACTGATGACATGGTGA 992  
 QY 299 pGlnAlaTyrPaspAlaGlnArgIlePheLysGluAla-GluLysPhePheValSerValG 319  
 |||||  
 Db 993 CCAAGGCTGAGATGCACAGATATATTCAGAGAGTCCGAGAACTTCTTGTATCTGTTG 1052  
 QY 319 lLeuProAsnMetThrGlnGlyPheTrp-GluAsnSerMetLeu-ThrAspProGlyAs 338  
 |||||  
 Db 1053 GTCTTCCTTATATAGACTATAGATTTCTGCGCAAAATTTCCATGCTATATACGAGCCAGAAA 1112  
 QY 338 nValGlnLysAlaValCysHis-ProThrAlaTrpAspLeuGlnLysGlyAspPheArg- 357  
 |||||  
 Db 1113 TGTTCAGAAAGCACTCTGCATCCCAAGCTTGGGACCTGGGAGGCGGACTTCAGAG 1172  
 QY 358 IleLeuMetCysThrLys-ValIleMetAspAspPheLeuThrAlaHisGluMetAl 377  
 |||||  
 Db 1173 ATCTTATGTGACAAAGGATACAAATGAGACACTTCTCTGACAGCTCATCATGAGATGGG 1232  
 QY 377 yHisIleGlnTyrAspMetAlaTyrAlaAlaGlnPro-PheLeuLeuArg-AsnGlyAla 396

Db 1233 GCATATCCACTATGATATGCGCATATATGCGGCCCAACCTTTTCTGTAGAAATGAGGCT 1292  
 QY 397 -AsnGlnGlyPheHisGlnAlaValGlyIleLeuMetSerLeuSerAlaIleThrProLy 416  
 |||||  
 Db 1293 TAAATGAGATTCATGCAATGCACTGTGGGAAATCATGCTCACTTGTGCGCCACACTTAA 1352  
 QY 416 SHISLeuLysSerIleGlyLeuLeuSerProAspPheGln---GluAspAsnGluThrG 435  
 |||||  
 Db 1353 GCATTTAAATTCATTTGCTTCTGTCTCACCAGGTTTTCACGACAGACATGAACACAA 1412  
 QY 435 uIleAsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMe 455  
 |||||  
 Db 1413 AATTAACCTCTGCTCAAAACAGACACACAGCATTTGTGGAGCTCCCATTTACTTACAT 1472  
 QY 455 tLeuGlnLysTrpArgTrpThrValPheLys-GlyGlnIleProLysAspGlnTrpMetL 475  
 |||||  
 Db 1473 GTTAGAAGATGGAGGTGGATGGCTTTTAAAGCGGGAAATTTCCCAAGACCAAGTGGGTGA 1532  
 QY 475 ySLys-TrpTrpGlnMetLysArgGlu-IleValGlyVal-ValGluProValProHisA 494  
 |||||  
 Db 1533 AAAAGTGTGGGAGATGAGAACCGAAGAAATAGTTGGGTGTGTGGAACCTGTGCCCCATG 1592  
 QY 494 srgIuThrTyr-CysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIle 513  
 |||||  
 Db 1593 ATGAACATATCTGTGACCCCGCATCTCTGTCCATGTTTCTAATGATTAATCATTCAT 1652  
 QY 514 ArgTyrThrArgThrLeu-TyrGlnPheGlnPheGln-GluAlaLeu-CysGlnAla 532  
 |||||  
 Db 1653 CGATTTATACACAGAGACCCGTGTTACCAATTCAGTTTCAAAAGAACACTTTGTCAAGCA 1712  
 QY 533 AlaLysHisGlnGlyProLeuHisLys-CysAspIle-SerAsnSerThrGlu---AlaG 551  
 |||||  
 Db 1713 GCTAAACATGAGGCGCTTGCACAAATGTGACATTTCTCAAAATTCACAGAAAGCTGCTG 1772  
 QY 551 yGlnLys-LeuPheAsnMetLeuArgLeuGlyLys-SerGlnProThrThrLeuAla 570  
 |||||  
 Db 1773 GACAGAACACTGTTCAATATGCTGAGGCTTGGAAACTGCAAAACCTGCGACCTTAGCAT 1832  
 QY 570 uGlnAsnValVal-GlyAlaLysAsnMetAsnValArgPro-LeuLeuAsnTyrPheGlu 589  
 |||||  
 Db 1833 GGAAATGTTTGTATAGACACCAAGAACATGATGAAGCCACCTGCTCAACTCTTGGAG 1892  
 QY 590 ProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGlyTrpSerThrAsp 609  
 |||||  
 Db 1893 CCTTATTTACTGCTGGAAGAACCAAGAACAAATCTTTTGGGATGAGATACCGGAC 1952  
 QY 610 TrpSerProTyrAlaAspGlnSerIle-LysValArgIleSerLeuLysSerAlaLeuG 629  
 |||||  
 Db 1953 TGGAGTCCATATGACAGACACAGCATCAAGTGAAGATGAAGCTTAATAATCAGCTTGG 2012  
 QY 629 y-AspLysAlaTyrGluTrpAsnAsp-AsnGlnMetTyrLeuPheArgSerSer-ValAl 648  
 |||||  
 Db 2013 CAGATTAAGCATATGATGAGACAGACCAATGAATGACCTGATCCGATCATCTGGTTGG 2072  
 QY 648 aTyrAla--MetArgGlnTyrPheLeu-LysValLysAsnGlnMetIleLeuPheGlyG 667  
 |||||  
 Db 2073 AATATTTAAATTTGAGGAGTACTTTTAAACAGTAAGTAATAATCGATATCTTTTGGGG 2132  
 QY 667 lGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheValThrA 687  
 |||||  
 Db 2133 AGGAGATGTGCGAGTGGCTAATTTGAAACCAAGAAATCTCTTAAATTTCTTGTACCTG 2192  
 QY 687 lAspLysAsnValSer-AspIleIleProArg-ThrGluValGluLysAlaIleArgMe 706  
 |||||  
 Db 2193 CACCTAAAAATGTGTCTGATATATCTCTTGAACACGAGTTGAAGGCAATCAGAT 2252  
 QY 706 tSerArgSerArg-IleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGlnPheLeuG 726  
 |||||  
 Db 2253 GTCCCGGAGCCGTACTCATGATGCTTCCCTGTAATGACAGACAGCTAGAGTTTCTGG 2312  
 QY 726 lYlIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTrpLeuIleValP 746  
 |||||

[illegible]



Db 781 TGYTNCNCNCACAYTNTNGNGATATGTGGGNGMNTTGTGACNAAYTNTAYCCN 840  
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 Db 841 YTNACGTCCTTTCGNCARARACCNAAATHGAGTACAGATGCATGATGAAATCAR 900  
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320  
 Db 901 GGNWGGAGYGCNGARWGNATHTTTCARCARGCNGARAAATTTTGTMSNGTNGCNTN 960  
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340  
 Db 961 CCNCAATATGACNCARCGNTTGTGGCNAAYGWSNATGTYTNACNGACCCNGCNAAYGNGN 1020  
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 Db 1021 AARGTGTGTGTCATCCACNCNCNTGGGAYTNGGNCAGVGNGATTTTGNATHTHARATG 1080  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisIleGluMetGlyHisIleGln 380  
 Db 1081 TGYACNAARGTACNATGAGAYAAATTTTNTACNGCCACAYCARATGGGNCATATGCAR 1140  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400  
 Db 1141 TAGGAYATGGCNTAYGCGNNGCARCCNTTGTNTYTNMGNAAYGNGCNAAYGARCGNTTY 1200  
 QY 401 HisGluAlaValGlyLysIleLeuMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
 Db 1201 CAYGAGCGTNGTNGNCARATHTATGWSNTYTNMSNGCNGCACCNACCNARATYTNARASN 1260  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440  
 Db 1261 ATHGAGVTYTNCCNMSNCAATTTTCARGARAYWSNGARACNARATHTAAATTTTNTYN 1320  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTrpMetLeuGluLysTrpArg 460  
 Db 1321 AACACAGCNYTNACNATHTGTNGNACNNTNCCNTTATATATATGTTNARARATGGGNCN 1380  
 QY 461 TrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrpGluMet 480  
 Db 1381 TGGATGTGTTTGTNGNGNCARATHTCCNARARCATGATGATGAARARATGGGAGARATG 1440  
 QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTrpCysAspPro 500  
 Db 1441 AARMGNGARATHTGTNGTNGTNGARCCNNTNCCNCAAGAYGARACNATAYGAYCCN 1500  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520  
 Db 1501 GCNMSNTYTNATTCATGTNMSNAAAYGAYTAYWSNTTATHTMGNTATATACNNGNACNATH 1560  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis 540  
 Db 1561 TATCATTTTCARTTTCARARARGCNTNTTTCARCGCGCAATATATAYAAAGNSNTNCAAY 1620  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560  
 Db 1621 AARTGAYATHTMSNAAAYSMACNGARCGNCGNCAAAATRYTNTNARATGTNMSNTYN 1680  
 QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsn 580  
 Db 1681 GGNAAAYWSNGARCCNTGACNNAARCGCNTNGARARAYGTGTGTGNGCNGCNGNNAATATGAY 1740  
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600  
 Db 1741 GTNAACGNTTNTNATAYATTTTCARCCNNTTNTTGTGATGTGTNARARARARAYGKN 1800  
 QY 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLeuVal 620  
 Db 1801 AAYWSNTTGTNGTNGAGAYACNGARGTGWSNCCNTATYGCNGAYCARMSNATHTHARGN 1860  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGlyTrpAsnAsnGluMet 640  
 Db 1861 MGNATHTMSNTNARMSNGCNTNTNGNCGNAAAYGCNTATGATGTGACNAAATAYAGARATG 1920

QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660  
 Db 1921 TTTYTNTTNGMWSNWSNGTNGCNTATGACNATGNGNARATRYTNTYMSNATHATHARAAV 1980  
 QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 Db 1981 CARCNGTNCNTTYYTNGNATGAGARGAGTNGTNGTMSNGAYTNTNARCCMNGCNTNWSN 2040  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700  
 Db 2041 TTTYATTTTYYTGTNACNWSNCCCARARATGTNMSNCAATGTNATHTCCMNGMWSNGARGTN 2100  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 Db 2101 GARGAYCGNATHMGNATGWSNMGNGMGNATHTAAGATNTTGTGANTYNAAYGAYAAV 2160  
 QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 Db 2161 WSNYTNGARRTYYTNGNATHCAYCCNACNTNGARCCNCCNTAYCARCCNCCNGTNACN 2220  
 QY 741 IleTrpLeuIleValPheGlyValAlaMetGlyValIleValValGlyIleValIleLeu 760  
 Db 2221 ATHTGGTNTATHATHTTGTGNGTGTNATGTCNTNTNGTNGTNGNATHTATHATHTYN 2280  
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780  
 Db 2281 ATHTGNCNCGNATHARAGMWSNARARARARARARARARARARARARARARARARARAR 2340  
 QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800  
 Db 2341 TAYGAYWSNATGAYATHTGNAARGNGARMSNAAAYCGNCGNTTTCARAAAYWSNGAY 2400  
 QY 801 ValGlnThrSerPhe 805  
 Db 2401 GCNCRACNMSNTTY 2415  
 Db 2401 GCNCRACNMSNTTY 2415  
 RESULT 17  
 AAQ10328  
 ID AAQ10328 standard; DNA; 2477 BP.  
 AC AAQ10328;  
 XX  
 AC  
 XX  
 DT 10-APR-1991 (first entry)  
 DT  
 XX  
 DE Encodes human testicular angiotensin conversion enzyme.  
 XX  
 KW human testicular angiotensin conversion enzyme; tACE;  
 KW male sterility; ss.  
 XX  
 OS Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 29..2227  
 FT /\*tag= a  
 FT /product= human tACE  
 FT Peptide 29..91  
 FT /\*tag= b  
 FT /label= signal peptide  
 XX  
 PM W09100354-A.  
 XX  
 PD 10-JAN-1991.  
 XX  
 PF 05-JUL-1990; 90WO-FR00513.  
 XX  
 PR 05-JUL-1989; 89FR-0009062.  
 XX  
 PA (INRM ) INST NAT SANTE RECH.  
 XX  
 PI Soubrier F, Albenc-Gelas F, Hubert C, Corvol P;  
 DR WPL: 1991-036748/05.



DR P-PSDB; AAR10426.

XX Nucleic acid - encoding human testicular angiotensin conversion  
 PT enzyme, used e.g. for in vitro detection of enzyme in organism  
 XX  
 PS Claim 1; Fig 1; 48pp; French.  
 XX  
 CC A bank of human testicular cDNA in Lambda gtl1 was screened with a  
 CC probe containing the final 3248 nucleotides of endothelial ACE. The  
 CC complete sequence of ACE was reconstructed from 4 separate clones.  
 CC It encodes a 711 amino acid mature protein and a 21 residue signal  
 CC peptide. The 228-2224 sequence is identical to the 1944-3940 region  
 CC of endothelial ACE. The isolated nucleic acid sequence is inserted  
 CC into a plasmid for expression of polypeptides. The invention also  
 CC covers parts of the sequence comprising all or part of the 29-229  
 CC sequence, any sequence differing from ACE only by silent  
 CC substitutions and nucleic acids which hybridise to ACE.  
 XX  
 SQ Sequence 2477 BP; 536 A; 811 C; 695 G; 435 T; 0 other;

Alignment Scores:

Pred. No.:	8.01e-121	Length:	2477
Score:	1344.00	Matches:	259
Percent Similarity:	60.97%	Conservative:	119
Best Local Similarity:	41.77%	Mismatches:	204
Query Match:	31.32%	Indels:	38
DB:	12	Gaps:	10

US-09-978-385-2 (1-805) x AAO10328 (1-2477)

QY 15 ThrAlaAlaGlnSer-----ThrIleGlnGlnAlaIleAlaSerThrPheLeuasp 30  
 |||||  
 DB 209 ACATAGAGCCAGAGCCCAAACTGCTGACTGATGAGCTGAGCCAGCAAGTTTGTGAG 266  
 QY 31 LysPheasnHISGluAlaGluaspLeuPheSerThrGlnSerSerLeuAlaSerThr 50  
 |||||  
 DB 269 GATATGACCGGACATCCAGAGGTGTGGAAGATGTCGAGCGAGCACTGGAACTAC 328  
 QY 51 AsnThrAsnIleThrGluGlu-----AsnValGlnAsnMet 62  
 |||||  
 DB 329 AACACCAACATCACCAGAGAGCCAGCAAGATTCGTCGACAGAGACATCAATGACC 388  
 QY 63 AsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnSerThrLeuAlaGlnMet 82  
 |||||  
 DB 369 AACCCACACC-----CTGAAGTACGGCACCCAGCCAGCCAGAG 424  
 QY 83 TyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeuGln 102  
 |||||  
 DB 425 TTTGATGTGAACCACTTCAGAACACCACTATCAAGCGCATCAAGAAAGGTTCCAGAC 484  
 QY 103 AsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeuAsnThr 122  
 |||||  
 DB 485 CTAGAACGGGAGCGCTCTGCTCCAGAGAGCTGAGAGGATCAACAGATCTCTTGAT 544  
 QY 123 MetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlnCysLeu 142  
 |||||  
 DB 545 ATGGAACACCACTACAGCGTGGCCACTGTGGCCACCCGAGATGCG-----AGCTGCGTG 598  
 QY 143 LeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGluArgLeu 162  
 |||||  
 DB 599 CAGCTCGAGCCAGATCTGAGCAATGTGATGCCACATCCCGAAATTTGAAGACCTTTA 658  
 QY 163 TrpAlaTrpGlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyrGluGlu 182  
 |||||  
 DB 659 TGGGATGGGAGGCTGGAGAGCAAGCGGGAGACCACTCTCCAGTTTACCCGAAA 718  
 QY 183 TyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAspTyrGlyAspTyr 202  
 |||||  
 DB 719 TACGTGAACATCAACACAGGCTGCCCGCTCAATGCTATGTAGTGAAGGAGGACTCG 778  
 QY 203 TrpArgLysAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGlyGlnLeu 222  
 |||||  
 DB 779 TGGAGGTCTATGTAGAGACCAACCACTCCCTGGAG----- 811

QY 223 IleGluAspValGluHisThrPheGlnGluIleLysProLeuTyrGlnHisLeuHisAla 242  
 |||||  
 DB 812 ---CAAGACCTGGAGCGGCTCTCCAGAGCTGCAGCCACTGACCACTGACATGCC 868  
 QY 243 TyrValArgAlaLysLeuMetAsnAlaTyr---ProSerTyrIleSerProIleGlyCys 261  
 |||||  
 DB 869 TACGTGGCCCGGCGCTGCACCTGACTACAGGGGCCAGACATCACTGAGAGGGCCC 928  
 QY 262 LeuProAlaHisLeuLeuGlyAspMetThrGlyArgThrPheThrAsnLeuTyrSerLeu 281  
 |||||  
 DB 929 ATTCTGCTACCTGCTGGGAGACATGTGGGCGAGCACTGGTCCAACTCTGACTGT 988  
 QY 282 ThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGlnAla 301  
 |||||  
 DB 989 GTGTGGCCCTTCCTTCAGGCCCTGATGAGACACCAAGAGGTATGTCAAGACAGGCC 1048  
 QY 302 TrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeuPro 321  
 |||||  
 DB 1049 TGGAGCCCGAGAGAGATTTAAAGAGCTGATGATTTCTTCACTCCGCGGGCTGCTG 1108  
 QY 322 AsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGlnLys 341  
 |||||  
 DB 1109 CCGTGGCTCTGAGTCTGAGAACACTGCTGAGAGAGCCACAGCGGGGGAG 1168  
 QY 342 AlaValLysHisProThrAlaTrpAspLeuGlyLysGly---AspPheArgIleLeuMet 360  
 |||||  
 DB 1169 GTGCTTGCACAGCGCTGGGCTGGACTCTACAAAGCAAGAGCTCCGATCAAGCAG 1228  
 QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisGlnMetGlnHisIleGln 380  
 |||||  
 DB 1229 TGCACCACTGAACTGGAGAGACTGTGTGTGGCCACCAAAATGGGCCATCCAG 1288  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPhe 400  
 |||||  
 DB 1289 TATTTCATGCAATCAAAAGACTTACCTGTGGCTTGAGGAGGAGGTCGCAACCCGGCTTC 1348  
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
 |||||  
 DB 1349 CATGAGGCCATTTGGGGGAGCGTACGCTCTGATGATGAGCCCAAGACCTGCAAGCT 1408  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGlnIleAsnPheLeu 440  
 |||||  
 DB 1409 CTCACCTGTGAGAGAGTAGGGGTGGAGAGGAC---GAGCATGACATCAACTTTCGATG 1465  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460  
 |||||  
 DB 1466 AAGATGCCCCCTGACAAGATCGCCCTTATCCCTTCAGCTACCTCGATCAGTGGCGC 1525  
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpArgLeuMet 480  
 |||||  
 DB 1526 TGGAGGTTATTTATGGAACATCAACAGAGAACTATTAACAGGAGAGTGGAGACCTTC 1585  
 QY 481 LysArgGluIleValGlyValValGluProValProHisAspGluThrTyrGlyAspPro 500  
 |||||  
 DB 1586 AGGCTGAAGTACAGAGGCGCTGCCCCAGTCCAGAGACATCAAGCTTATGACCCA 1645  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520  
 |||||  
 DB 1646 GGGGCGAAGTCCACATCTCTTACGCTGCTTACATAGGATCTTGTACGCTTCAATC 1705  
 QY 521 TyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysGlnLysProLeuHis 540  
 |||||  
 DB 1706 ATCCACTTCCAGTCCACAGAGGACTGTGCCAGGACGCTGCCACAGCGGCCCTGACC 1765  
 QY 541 LysCysAspLysSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560  
 |||||  
 DB 1766 AAGTTGTACATATCACAGTCCAAAGAGCGCGGAGCGCTGGCGACCCGATGAAGCTG 1825  
 QY 561 GlyLysSerGluProThrPheThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetLys 580  
 |||||  
 DB 1826 GGTCTTGAAGGCGGTGGCGGAGAACCAAGCATGATCACGGGCGAGCCCAACATGAGC 1885

QY 581 ValArgProLeuLeuAsnTyr-PheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600  
 Db 1886 GCCTGGCCATGTTAGCTACTTCAAGCCCTGCTGGAGCTGGCTCCGACGAGACAG 1945  
 QY 601 -----AsnSerPheValGlyTyr---SerThrAspTrpSerProTyrAlaAspGlnSer 617  
 Db 1946 CTGCATGGGGAGACAGCTGGGCGGACGAGTACACTGAGCGGACACTCGCTGCTCA 2005  
 RESULT 18  
 AAA38330  
 ID AAA38330 standard; DNA; 4020 BP.  
 XX  
 AC AAA38330;  
 XX  
 DT 21-AUG-2000 (first entry)  
 XX  
 DE Human angiotensin-converting enzyme (ACE) coding region.  
 XX  
 KW Angiotensin-converting enzyme gene; ACE; coding region; polymorphism;  
 KW polymorphic marker; cardiovascular disease; myocardial infarction;  
 KW unstable angina; hypertension; atherosclerosis; stroke; prognosis;  
 KW drug screening; treatment outcome; human; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200022166-A2.  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 13-OCT-1999; 99WO-IB01678.  
 XX  
 PR 14-OCT-1998; 98US-0104286.  
 PR 14-OCT-1998; 98US-0104302.  
 XX  
 PA (EURO-) EUROPA MEDICAL AB.  
 XX  
 PI Norberg LT, Andersson MK, Lindstrom PHR, Jonsson L;  
 XX  
 DR WPI: 2000-318010/27.  
 XX  
 DX  
 XX  
 PT Assessing cardiovascular status in humans involves comparing test  
 PT polymorphic pattern comprising polymorphic positions within genes  
 XX encoding specific proteins, with reference polymorphic pattern  
 XX  
 PS Disclosure: Page 114-115; 126pp; English.  
 CC The invention relates to a novel method of assessing the cardiovascular  
 CC status in an individual and to newly identified polymorphisms in the  
 CC genes encoding angiotensin-converting enzyme (ACE), angiotensin II  
 CC receptor type 1 (AT1) and type 2 (AT2), angiotensinogen (AGT), renin,  
 CC aldosterone synthase, endothelin receptor type A and beta-adrenergic  
 CC receptors 1 and 2. The method comprises determining the sequence at one  
 CC or more polymorphic positions within these genes, and comparing the  
 CC pattern of polymorphisms from the individual with a reference polymorphic  
 CC pattern obtained from a population of individuals exhibiting a  
 CC predetermined cardiovascular disease status. The polymorphic markers are  
 CC useful for determining the predisposition of an individual to  
 CC cardiovascular disorders such as myocardial infarction, unstable angina,  
 CC hypertension, atherosclerosis and stroke. They are also useful for  
 CC predicting the likely cardiovascular status of a patient given a  
 CC treatment regimen comprising administration of cardiovascular drugs  
 CC (e.g., ACE inhibitors, beta-adrenergic receptor antagonists (beta-  
 CC blockers) or calcium channel blockers). One or more polymorphic markers  
 CC provides a basis for predicting the outcome of a treatment regimen.  
 CC Fragments of the genes comprising a polymorphic site may be used as  
 CC primers and probes for detecting genetic polymorphisms or in molecular  
 CC library arrays for high throughput screening. The genes, and the proteins  
 CC they encode are useful in the screening of potential cardiovascular  
 CC drugs. Determination of an individual's polymorphic pattern reduces or  
 CC eliminates trial and error in selecting a treatment for a particular  
 CC individual cardiovascular patient. It also provides the ability to  
 CC eliminate patients from clinical trials who are predicted to be  
 CC non-responsive, or at a risk for an adverse response, to a particular

CC treatment regimen. Adverse results in an early trial can be evaluated to  
 CC identify polymorphic patterns so that the adverse results can be  
 CC correlated with a sub-population of the test population, permitting  
 CC exclusion of such sub-populations from the treatment group. Beneficial  
 CC drugs can be approved for use in the appropriate population, thereby  
 CC decreasing the number of patients required for a clinical trial, which in  
 CC turn decreases the duration and cost of such trials. Sequences A38328 and  
 CC A38330 represent, respectively, intron 16 and the coding region of  
 CC the human ACE gene (GenBank X52855, J04144). The polymorphic sites  
 CC identified are 375A/C, 582C/T, 731A/G, 1060G/A, 1215C/T, 2193G/A,  
 CC 2328A/G, 2741G/T, 3132C/T, 3387T/C, 3503G/C, 3906G/A; and a deletion of  
 CC nucleotides 1451-1783 in intron 16.  
 CC  
 XX  
 SQ Sequence 4020 BP; 857 A; 1261 C; 1174 G; 728 T; 0 other;

## Alignment Scores:

Alignment No.	Score	Length	Matches	Conservative	Mismatches	Indels	Gaps
1	7.73e-120	4020	1337.00	235	118	34	9
2	Percent Similarity:	61.05%	Conservative:	204			
3	Best Local Similarity:	41.73%	Mismatches:	204			
4	Query Match:	31.16%	Indels:	34			
5	DB:	21	Gaps:	9			

US-09-978-385-2 (1-805) x AAA38330 (1-4020)

QY 20 ThrIleGluGluGlnAlaLysTyrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 39  
 Db 1952 ACTGATGAGCGGTGAGCGCAGCAGGAGTTGTGAGGAATATGACCGGACATCCAGGAGTG 2011  
 QY 40 PheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGlu----- 57  
 Db 2012 TCGAACAGATATGCGCGAGCACTGACACTACAAACCCACCACTCCACAGACAGCAGC 2071  
 QY 58 -----AsnValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAla 71  
 Db 2072 AAGATTCTGCTGCAGAAACATGCAATATGCCAACACAC----- 2113  
 QY 72 PheLeuLysGluGlnSerTrpLeuAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeu 91  
 Db 2114 -----CTGAAGTACGGCGACCCAGCCAGGAGTGTGATGTGAAACCGTGCAGAACACC 2167  
 QY 92 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnLysSerValLeuSerGluAsp 111  
 Db 2168 ACTATCAAGCGCATCAATAAAGAGTTACAGCTAGACCTAGACGAGCGCTGCTGCCAG 2227  
 QY 112 LysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLys 131  
 Db 2228 GAGCTGAGGAGTACACACATCTCTGTGATATGGAACACACCTACAGCGGCCACT 2287  
 QY 132 ValCysAsnProAspAsnProGlnGluCysLeuLeuGluProGlyLeuAsnGluIle 151  
 Db 2288 GTGTGCCACCCGAAATGCG-----AGCTGCTGCAGCTGACGCCAGATCTGACGAATG 2341  
 QY 152 MetLysAsnSerLeuAspTyrAsnGluArgLeuTrpAlaTrpGluSerTrpArgSerGlu 171  
 Db 2342 ATGGCCACATCCCGGAATATGAAGACCTGTTATGGCATGAGGAGCTGGCAGACAG 2401  
 QY 172 ValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGluMetAla 191  
 Db 2402 GCGGGAGAGGCCATCTCCATCTTTACCCGAATATCTGGAACATCATACACAGGCTGCC 2461  
 QY 192 ArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGly 211  
 Db 2462 CGGCTCAATGCTATGTAGATGAGCGGAGACTCGTGAAGCTATATGACGAGACACATCC 2521  
 QY 212 ValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGlnHisThrPheGlu 231  
 Db 2522 CTGGAG-----CAAGACCTGGAGCGGCTTTCAG 2551  
 QY 232 GluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAla 251  
 Db 2552 GAGCTGCAGGACATCTACCTGATACCTGATGTCAGCGCGCGGCGCTGCACCGCTCAC 2611

```

QY 252 Tyr---ProSerTyrIleSerProIleGlyCysLeuProAlaHisIleLeuGlyAspMet 270
DB 2612 TACGGGCCCCAGACACCTGAGGCCCATCTTCCCTGACCTCTGGGGAGACATG 2671
QY 271 TrpGlyArgPheThrPheAsnLeuTyrSerLeuThrValIleProPheGlyGlnLysProAsn 290
DB 2672 TGGGGCCAGACCTGGTCAACATCTATGATCTGGTGGCTCCCTTCACGCCCTCCG 2731
QY 291 IleAspValThrAspAlaMetValAspGlnAlaTrpAspAlaGlnArgIlePheLysGlu 310
DB 2732 ATGACACACACAGAGCCTATGCTAAAGCAGCGCTGGACGCCAGAGATCTTTAAAGAG 2791
QY 311 AlaGlyLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTrpGluAsn 330
DB 2792 GCTGATGATTTCTTCCACTCCCTCGGGGCTGCTGCCCGCTCCCTGAGTCTTGAAACAG 2851
QY 331 SerMetLeuThrAspProGlyAsnValGlnLysAlaValCysHisProThrAlaTrpAsp 350
DB 2852 TCGATGCTGGAGAACCCAGCCAGCGGGAGGTGGTGTCTGTCACGCCCTCGCTGGGAC 2911
QY 351 LeuGlyLysGly---AspPheArgIleLeuMetCysThrLysValThrMetAspAspPhe 369
DB 2912 TTCTACACAGCGCAGAGACTCTCCGGATCAAGCAGTCAACCCAGCTGAGACTGGAGACTG 2971
QY 370 LeuThrAlaHisHisGlyMetGlyHisIleGlnTyrAspMetAlaTyrAlaAlaGlnPro 389
DB 2972 GTGTGGGCCCCACCAATGGCCACATCCAGTATTTCAAGTACCAAGACTACTACT 3031
QY 390 PheLeuLeuArgAsnGlyAlaAsnGlyLeuPheHisGlyAlaValGlyGlnIleMetSer 409
DB 3032 GTGGCTTGAGAGGAGGAGGAGGCCAACCCCGGCTTCCATGAGCCATGGGAGGTCTAGCC 3091
QY 410 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 429
DB 3092 CTCTCAGTGTCTAGCCCAACACCTGACAGCTCAACCCGCTGACAGAGGAGGAGG 3151
QY 430 GluAspAsnGlyThrGlnIleAsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThr 449
DB 3152 AGCGAC---GAGCATGACATCACTTCTGTATGAAGATGGCGTCAACAAGATGCCCTT 3208
QY 450 LeuProPheThrTyrMetLeuGlyLysTrpArgTrpMetValPheLysGlyGlnIlePro 469
DB 3209 ATCCCTTCACCTACCTCGTGCATCAGTGGCGCTGGAGGATTTGATGCAAGCATACCC 3268
QY 470 LysAspGlnTrpMetLysLysTrpTrpGlnMetLysArgGlnIleValGlyValGlu 489
DB 3269 AAGGAGACTATTAACACAGAGGTGGTGGAGCTCAGCTGAAGTACAGAGGCCCTGCCCC 3328
QY 490 ProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAsp 509
DB 3329 CCAAGTCCCAAGACTCAAGGTGACTTTGACCCAGGGGCCAAGTTCACATCTCTTCTAGC 3388
QY 510 TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnGlnAlaLeu 529
DB 3389 GTGGCTTACATCAGTACTTGTTCAGCTTATCATCATCAGTTCCAGTTCCAGAGGACTG 3448
QY 530 CysGlnAlaAlaLysHisGlyGlyProLeuHisLysCysAspIleSerAsnSerThrGln 549
DB 3449 TGGCAGGACACTGGCCACAGCGGCCCTGCACAAAGTGCATCTACACAGTCCCAAGAG 3508
QY 550 AlaGlyLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProThrPheAla 569
DB 3509 GCGGGGAGAGGCGCTGGCGACCCCAATCAAGTGGCTTCTGATGGCCCGCGAAGCC 3568
QY 570 LeuGlnAsnValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu 589
DB 3569 ATGACGCTGATCAAGGCGCCAGCCCAACATAGAGGCGCTCGGCAATGTTGACTCTTCAAG 3628
QY 590 ProLeuPheThrTrpLeuLysAspGlnAsnLys-----AsnSerPheValGlyTrp-- 606
DB 3629 CCGGTGCTGAGTGGCTGCGCCAGGAGAACAGCTGATGAGGAGAAAGCTGGCTGGCG 3688
QY 607 SerThrAspTrpSerProTyrAlaAspGlnSer 617

```

```

DB 3689 CAGTACACTGGACGCGCAACTCCGCTGCTCA 3721
RESULT 19
AAK93395
ID AAK93395 standard; DNA: 4022 BP.
XX
AC AAK93395;
XX
DE 27-JUN-2002 (first entry)
XX
DE DNA of APP related human homologue hcp51674.
XX
KW Neuroprotective; nootropic; transgenic fly; Alzheimer's disease; Abeta;
KW amyloid precursor protein; tissue-specific expression control; human APP;
KW APP pathway modulator; gene therapy; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 148..3945
FT /tag="a
FT /product="Protein of human homologue hcp51674"
FT /note="No start codon"
XX
MO200226820-A2.
XX
PD 04-APR-2002.
XX
PF 01-OCT-2001; 2001WO-BP1345.
XX
PR 29-SEP-2000; 2000US-236893P.
XX
PR 14-JUN-2001; 2001US-298309P.
XX
PA (NOVS ) NOVARTIS AG.
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI Cohen D, Dengler UJ, Finelli AL, Freuler F, Konsolaki M;
PI Reinhardt MWM, Zusan S;
XX
DR WPI: 2002-315796/35.
DR P-PSDB; AAO20501.
XX
PT New transgenic fly, containing DNA encoding an Abeta portion of human
PT APP, useful for identifying agents which modulate the APP pathway and
PT which can be used to treat Alzheimer's disease -
XX
PS Example 4; Page 93-94; 129pp; English.
XX
CC The invention relates to a transgenic fly whose genome comprises DNA
CC encoding a polypeptide having the Abeta portion of human amyloid
CC precursor protein (APP), fused to a signal sequence. The DNA sequence
CC encodes a 123 (Abeta40) or 129 (Abeta42) amino acid sequence, given in
CC the specification. The DNA sequence is operably linked to a tissue-
CC specific expression control sequence. Expression of the sequence gives
CC the fly an altered phenotype. The purpose of the invention is for
CC identifying agents that inhibit or promote the expression and/or function
CC of genes or encoded polypeptides which modify the APP pathway. The agent
CC is a compound, triple helix DNA, antisense oligonucleotide, double
CC stranded RNA molecule, ribozyme, or particularly an antibody. It is used
CC to treat conditions such as Alzheimer's disease. The agent can be used as
CC an APP pathway modulator or in gene therapy. This polynucleotide sequence
CC represents the DNA of the APP related human homologue hcp51674.
XX
SQ Sequence 4022 BP; 856 A; 1262 C; 1174 G; 728 T; 2 other:
Alignment Scores:
Pred. No.: 7.74e-120 Length: 4022
Score: 1337.00 Matches: 255
Percent: 61.05% Conservative: 118
Best Local Similarity: 41.73% Mismatches: 204
Query Match: 31.16% Indels: 34
DB: Gaps: 24

```

US-09-978-385-2 (1-805) x AA09395 (1-4022)

QY 20 ThrIleGluGlnAlaIleuThrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 39  
 Db 1954 ACTGATGAGGCGCCAGCAAGTTTCTGACGCAATATGACCGGACATCTCCAGTGGTG 2013  
 QY 40 PheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGlu----- 57  
 Db 2014 TGGAAAGCATATGCGAGGCGCACTGAACTTACACACCAACATATCACCAGACAGACACC 2073  
 QY 58 -----AsnValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAla 71  
 Db 2074 AAGATTCTGCTGCAGACAGACATGCAAATATGCAACACACACC----- 2115  
 QY 72 PheLeuLysGluGlnSerThrIleuAlaGlnMetTyrTrpLeuGlnGluIleGlnAsnLeu 91  
 Db 2116 -----CTGAAGTACGGGACCCAGGCGCAAGTTTGTATGTGAACCAAGTTCAGAACACC 2169  
 QY 92 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAsp 111  
 Db 2170 ACTATCAAGCGGATATTAAGAGTTGACAGACCTAGAACGGGCGACGCTGCTGCCAG 2229  
 QY 112 LysSerLysAlaGlyLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLys 131  
 Db 2230 GAGCTGAGAGAGTACACAAAGATCCTGTGGATGAAACACACCTACAGCGTGGCCACT 2289  
 QY 132 ValLysAsnProAspAsnProGlnIleLysLeuLeuGluProGlyLeuAsnGluIle 151  
 Db 2290 GGTGTCCACCCGAAATGCG-----AGTGTCCGTCAGCTGACCGCATGATGACGANTGTG 2343  
 QY 152 MetAlaAsnSerLeuAspTyrAsnGlnLysLeuTrpAlaTrpGlySerTrpArgSerGlu 171  
 Db 2344 ATGGGACAGTCCCGGAATATGAGACCTGTATGGCATGGGAGGCGTGGGAGACACAC 2403  
 QY 172 ValGlyLysGlnLeuAspProLeuTyrGluGlyValValLeuLysAsnGluMetAla 191  
 Db 2404 GCGGGAGAGGCGCTCCAGTTTACCGGAATAGCGTGAACCTCATACACCGGCTCC 2463  
 QY 192 ArgAlaAsnHisTyrGlnLysTyrGlyLysAspTyrTrpArgGlyAspTyrGluValAsnGly 211  
 Db 2464 CCGGTCAATGGCTATGTAGATGACAGGGAGCTGTGGAGTGTATGACAGACACACATCC 2523  
 QY 212 ValAspGlyTyrAspTyrSerAlaGlyLeuIleGlyLeuAspValGlnHisThrPheGlu 231  
 Db 2524 CTGGAG-----CAAGACCTGAGACGGGCTTCCAG 2553  
 QY 232 GluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAla 251  
 Db 2554 GAGCTGACAGCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2613  
 QY 252 Tyr-----ProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuGlyAspMet 270  
 Db 2614 TACGGGCGCCAGACATCACTGAGGGGCCCATTTCCGTCACCTGTGGGGAACATG 2673  
 QY 271 TrpGlyArgPheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 290  
 Db 2674 TGGGCGCAGACCTGTCAACATCTATGACTGTGTGTGCTCCCTTCCATCAACCCCTCG 2733  
 QY 291 IleAspValThrAspAlaMetValAspGlnAlaTrpAspAlaArgIlePheLeuGlu 310  
 Db 2734 ATGGACACCAAGAGGCTATGCTAAAGCAGGCTGAGACCCCAAGAGAGATTTAAGAG 2793  
 QY 311 AlaGlyLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTrpGluAsn 330  
 Db 2794 GCGTATGATTTCTCACTCCCTGGGGCGTGTGCGCGTCCCTCGAGTTCGGAACAG 2853  
 QY 331 SerMetLeuThrAspProGlyAsnValGlyLysAlaValCysHisProThrAlaTrpAsp 350  
 Db 2854 TCGATCTGAGAGACCAACGAGCGGAGGTGTGTGCTGCACAGCCCTGCGGCG 2913  
 QY 351 LeuGlyLysGly-----AspPheArgIleLeuMetCysThrLysValThrMetAspPhe 369

Db 2914 TTCTAACAGCGCAAGGACTCCCGATCAAGCAATGACACCGCTGAATCTGGAGGACTTG 2973  
 QY 370 LeuThrAlaHisIleGluMetGlyHisIleGlnTyrAspMetAlaThrAlaIleGlnPro 389  
 Db 2974 GTGGGCGCCACACAGAAATGGGCGACATCCAACTATTTCTGACATGACAAAGACTTACT 3033  
 QY 390 PheLeuLeuArgAsnGlnAlaAsnGluGlyPheHisGluAlaValGlyLuiLeuMetSer 409  
 Db 3034 GTGGGCTTGAAGGAGGGGTGCACACCGGCTTCATGAGGAGCCATTGGGAGCGTGAAGC 3093  
 QY 410 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 429  
 Db 3094 CTCTCACTGTCTAGCGCCAAAGCTGCACAGACTCACTCACTCACTCACTCACTCACTCA 3153  
 QY 430 GluAspAsnGluThrGluIleAsnPheLeuLeuGlnAlaLeuThrIleValGlyThr 449  
 Db 3154 AGCGAC---GAGCATGACATCAACTTTCTGATTAAGATGGCCCTTACAAAGATCGCTTT 3210  
 QY 450 LeuProPheThrTyrMetLeuGluLysTyrPargTrpMetValPheLysGlyGluIlePro 469  
 Db 3211 ATCCCTTCAAGTACCTCGTCGATGACAGTGGCGCTGAGGAGTATTTGATGAAACATCAC 3270  
 QY 470 LysAspGlnTrpMetLysLysTrpTrpGluMetLysArgGluIleValGlyValGlu 489  
 Db 3271 AAGAGAGACTATTAACCGAGGTGTGTGAGCGCTGAGCTGAAGTACCAAGGCGCTGCC 3330  
 QY 490 ProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAsp 509  
 Db 3331 CCAATGCCGACGACTCAAGTGAAGTGTGACACCGAGGCGCAAGTTCACATTCCTTAC 3390  
 QY 510 TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnGluAlaLeu 529  
 Db 3391 GTGGCTTACATCAGGTCTTGTGACGCTTCATCATCTCAAGTTCACAGTTCACAGGCGACTG 3450  
 QY 530 CysGlnAlaAlaLysHisGluGluProLeuHisLysCysAspIleSerAsnSerThrGlu 549  
 Db 3451 TGGCAGGAGCTGGGCGCACAGGGCGCCCTCACAAGTGTATCTTACATCCAGTCAAGAG 3510  
 QY 550 AlaGlyLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProThrPheLeuAla 569  
 Db 3511 GCGCGGAGGCGCTGGGCGACCGCCATGAACTGGGCTTCAGTGAAGCGCGGCGGAGGCC 3570  
 QY 570 LeuGluAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu 589  
 Db 3571 ATGAGCTGATCAGGCGCGACCCACATGAGCGCTGGCCATGTTGAGACTTTCAG 3630  
 QY 590 ProLeuPheThrTrpLeuLysAspGlnAsnLys-----AsnSerPheValGlyTrp--- 606  
 Db 3631 CCGCTGTGAGCTGCTGCCCGACGAGAGACAGCTGATGGGAGAAAGCTGGGCTGGCG 3690  
 QY 607 SerThrAspTrpSerProTyrAlaAspGlnSer 617  
 Db 3691 CAGTACAACGTGAGCGCAACTCGGCTCGCTCA 3723

RESULT 20  
 AA004027  
 ID AA004027 standard; cDNA; 4024 BP.  
 AC AA004027;  
 XX  
 DT 07-SEP-1990 (first entry)  
 XX  
 DE human angiotensin converting enzyme (ACE) gene.  
 XX  
 KW human angiotensin converting enzyme; hypertension;  
 KW bradykinin;  
 XX  
 OS synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT sig\_peptide 23..51  
 FT /\*tag= a  
 FT CDS 23..3944

```

*tag= b
XX
XX WO003435-A.
XX
XX 05-APR-1990.
XX
XX 29-SEP-1989; 89WO-FR00469.
XX
XX 27-SEP-1988; 88FR-0012620.
XX
XX (INNM ) INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE.
XX
XX Soubrier F, Alhenc-Gelas F, Hubert C, Corvol P;
XX WPI; 1990-058128/08.
XX DR P-PSDB; AAR04111.
XX
XX New DNA encoding human angiotensin converting enzyme used eg in
XX diagnosis of hypertension, evaluation of enzyme inhibitors
XX
XX Disclosure; P; French.
XX
XX This sequence encodes pre-ACE. In the mature protein the 29 amino acid
XX signal peptide is absent. ACE hydrolyses angiotensin I and/or kinins,
XX most notably bradykinin. Vectors contg. the given sequence are used to
XX produce ACE for therapeutic use eg to hydrolyse kinins implicated in
XX inflammation.
XX
XX Sequence 4024 BP: 857 A; 1263 C; 1175 G; 729 T; 0 other:
XX
Alignment Scores:
Pred. No.: 7.74e-120 Length: 4024
Score: 1337.00 Matches: 255
Percent Similarity: 61.05% Conservative: 118
Best Local Similarity: 41.73% Mismatches: 204
Query Match: 31.16% Indels: 34
DB: 11 Gaps: 9
US-09-978-385-2 (1-805) x AAQ04027 (1-4024)
QY 20 ThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 39
DB 1932 ACTGATGAGGCTGAGGAGCCAGCAAGTTTGTGAGGAATGTACCCGACATCCAGGTG 2011
QY 40 PheTyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnIleThrGlu----- 57
DB 2012 TGGACGAGTATGCCGAGGCCACCTGGAAGTCACTACCAACCAATCCACAGAGCAGC 2071
QY 58 -----AsnValGlnAsnMetAsnAlaGlyAspLysTyrSerAla 71
DB 2072 AAGATTCTGCTGCAGAAAGACATGCAATAGCCAAACCCACC----- 2113
QY 72 PheLeuLysGlnGlnSerThrLeuAlaGlnMetTyrProLeuGlnGlnIleGlnAsnLeu 91
DB 2114 -----CTGAAGTACGGCAGCCAGGAGGAAATTGATGGAACCAAGTTGCCAACAAC 2167
QY 92 ThrValLysLeuGlnLeuAlaLeuGlnLysAsnGlnSerValLeuSerGluAsp 111
DB 2168 ACTATCAGCGCATCAATAAAGATTTCAGAGCTAGAACGGGAGCGCTCCGCCAG 2227
QY 112 LysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLys 131
DB 2228 GAGCTGGAGGAGTACAAAGATCTCTGTGATGTGAACCACTACAGCGTGGCCACT 2287
QY 132 ValCysAsnProAspAsnProGlnGluCysLeuLeuGlnProGlyLeuAsnGlnIle 151
DB 2288 GTGTGCCACCGCATGGC-----AGCTGCTGCAGCTCGAGCCAGATCTGACGAATGTG 2341
QY 152 MetAlaAsnSerLeuAspTyrAsnGlnLysArgLeuThrPalaArgLysTyrArgSerGlu 171
DB 2342 ATGGCCACATCCCGGAATATGAGACCTGTTATGGCATGGAGGGCTGGCGAGAACAG 2401
QY 172 ValGlyLysGlnLeuArgProLeuTyrGlnGluLysValIleLeuLysAsnGlnMetAla 191

```

```

DB 2402 GCGGGAGAGCCATCCCTCACTTTTACCCGAAATACGTGAACTCATCAACAGGCTGCC 2461
QY 192 ArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrpArgLysPyrGluValAsnGly 211
DB 2462 CGGCTCAATGCGTATGTAGATGACAGGGAGACTCGTGGAGGCTATGTACGAGACCACTCC 2521
QY 212 ValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGlnIleThrPheGlu 231
DB 2522 CTGGAG-----CAAGACCTGGAGCGGCTTCCAG 2551
QY 232 GluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAla 251
DB 2552 GAGCTGACAGCCTCTACCTCACTCAATGCTGAGCGGCGCCCTGCACCGCTCAC 2611
QY 252 Tyr--ProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuGlyAspMet 270
DB 2612 TACGGGCGCCAGCAGATCAACCTGAGGGGCCATTCTCTGCACTCGGGGAGCAATG 2671
QY 271 TrpGlyArgPheTrpTrpAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 290
DB 2672 TGGCGGAGACCTGTCCAACTATGACTGTGGTGGCCCTTCCTCAGCCCTCG 2731
QY 291 IleAspValThrAspAlaMetValAspGlnAlaThrAspAlaGlnArgIlePheLysGlu 310
DB 2732 ATGACACACACAGAGGCTATGCTAAAGCAGGCTGAGCGCCAGGAGAGATTAAAGAG 2791
QY 311 AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnLysPheTrpGluAsn 330
DB 2792 GCTGATGATTTCTTCACTCCCTCGGGCTGCTGCGCCCTGCTGCTGCTGGAACAAG 2851
QY 331 SerMetLeuThrAspProGlyAsnValGlnLysAlaValCysHisProThrAlaTrpAsp 350
DB 2852 TCGATGTGAGAGAACCCAAACAGCGGCGGAGGTGTCTGCCACCCCTCGCGTGGAG 2911
QY 351 LeuGlyLysGly--AspPheArgIleLeuMetCysThrTyrLysValThrMetAspAspPhe 369
DB 2912 TTTCACAAAGGCAAGGACTTCGCGATCAAGCACTGACCAACCTGAACTGGAGGACTG 2971
QY 370 LeuThrAlaHisHisGluMetGlyHisIleGlnTyrAspMetAlaTyrAlaAlaGlnPro 389
DB 2972 GTGGTGGCCACACAGCAAAATGGGCCACATGCTGATATTCAGATGACAAAGACTTACT 3031
QY 390 PheLeuLeuArgAsnGlnAlaAsnGlnGlyPheHisGluAlaValGlyLysIleMetSer 409
DB 3032 GTGGCTTGGAGGAGGCTGCCAACCCCGCTTCCATGAGGCGCAATTGGAGAGCTGACG 3091
QY 410 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 429
DB 3092 CTCTCAATGCTACAGCCCAAGACACCTGCACAGTCTCAACTGCTGACAGTGAAGGTG 3151
QY 430 GluAspAsnGluThrGlnIleAsnPheLeuLeuLysGlnAlaLeuThrIleValGlyThr 449
DB 3152 AGCGAC--GAGCATGATCAACTTCTGATGATGAAATGATGCCCTTGACAAATGCCCTTT 3208
QY 450 LeuProPheThrTyrMetLeuGlnLysTyrPargTrpMetValPheLysGlyGlnIlePro 469
DB 3209 ATCCCTTCAGCTACCTCGTCAATGAGTGGCTGAGGATTTGATGTAAGACATCAC 3268
QY 470 LysAspGlnTrpMetLysTyrTrpGluMetLysArgGlnIleValIleValIleValGlu 489
DB 3269 AAGGAGAACTATTAACAGAGGAGTGTGAGAGCTCAGGCTAAAGTACAGAGGCTCTG 3328
QY 490 ProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSerAsnAsp 509
DB 3329 CCAATGCCCAAGGACTCAAGTGTGATTCACCAAGGAGGAAATGCTTCTTCAAGC 3388
QY 510 TyrSerPheIleArgTyrTyrThrArgTrpLeuTyrGlnPheGlnPheGlnIleVal 529
DB 3389 GTGCTTACATCGAGTACTTGTGACCTTCATCATCCAACTTCAAGTTCACAGGAGCAG 3448
QY 530 CysGlnAlaAlaLysHisGlnGluProLeuHisLysCysAspIleSerAsnSerThrGlu 549

```

```
Db 3449 TGCCAGGACGCTGGCCACAGCGGCCCTCGACAAAGTGTGACATCTACCAAGTCCAGAGAG 3508
QY 550 AlagIyGInLysLeuPheAsnMetLeuArgLeuGlyLysSerGIuProTfPThrLeuAla 569
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 3509 GCCGGGCGAGCGCGCTGGCGACCGCCATGAAGCTGGGCTTCAGTAGCCGCGCGGAAGCC 3568
QY 570 LeuGIuAsnValIAlAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGIu 589
    ::::: ||| |||||::: |||||::: |||||::: |||||::: |||||:::
Db 3569 ATGCAGCTGATCAGCGGCCACGCCACATGAGCGCGCTCGGCCATGTTGAGCTACTTCAG 3628
QY 590 ProLeuPheThrTfPLeuLysAspGlnAsnLys-----AsnSerPheValGIyTfP--- 606
    ||||| |||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
Db 3629 CCGCTGCTGAGACTGCTCCGACAGGAGAACGAGCTGATGGGAGAGAGCTGGGCTGGCGG 3688
QY 607 SerThrAspTfPSerProTfAlaAspGlnSer 617
    ::|||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db 3689 CAGTACAACGTGAGCGCCGAACTCCGCTCGCTCA 3721
```

Search completed: March 7, 2003, 07:47:45  
Job time : 494 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 7, 2003, 07:29:14 ; Search time 2368 seconds

(without alignments)  
5505.645 Million cell updates/sec

Title: US-09-978-385-2

Perfect score: 4291

Sequence: 1 MSSSSWLLSLVAVTAASST.....ISKGNPNPGFNTDVTQTSF 805

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/gen2.1/USPTO.spool/US09978385/rnat\_28022003\_104716\_1351/app\_query.fasta\_1.967  
-DB=EST -QWRT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=20 -MODE=LOCAL  
-OUTPM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09978385.0CGN.1.1.2874.0rnat.28022003\_104716\_1351 -NCP=6 -ICPU=3  
-NO\_XLPYX -NO\_MMAP -LARGOQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120  
-WART-TIMEOUT=40 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlm:\*  
5: em\_estloy:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mem:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_oth:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4249	99.0	3337	11	BC032938
2	1452	33.8	987	9	AL551235
3	1071	25.0	1603	11	AK008530
4	1054.5	24.6	763	13	BI913504
5	1048.5	24.4	879	12	BG401683
6	1031	23.0	766	13	BG62298
7	1022	22.8	676	13	BI561069
8	979	22.8	635	12	BG722079
9	901	21.0	609	13	BI561359
10	864	20.1	965	12	BF789159
11	857	20.0	555	9	AI162058
12	819	19.1	686	10	BB652968
13	781	18.2	452	12	BG428060
14	749.5	17.5	556	13	BM030353
15	730.5	17.0	741	12	BG772384
16	687.5	16.0	800	13	BI926471
17	684	15.9	1154	13	BM460886
18	680	15.8	900	13	BI415293
19	673	15.7	1018	13	BM560537
20	672.5	15.7	778	13	BI831019
21	661.5	15.4	848	13	BI904144
22	659.5	15.4	842	12	BF784147
23	646	15.1	848	13	BI766340
24	643	15.0	492	10	AM260204
25	625	14.6	738	9	AL039738
26	622	14.5	343	13	BI830936
27	615	14.3	887	10	BI196673
28	614	14.3	347	13	AM848579
29	593	13.8	666	13	BI686319
30	588.5	13.7	725	13	BI821905
31	578	13.5	751	13	BI828973
32	575.5	13.4	471	10	AM358360
33	564.5	13.2	823	13	BI914635
34	564	13.1	606	10	AM322773
35	553	12.9	728	10	AM026259
36	549	12.8	459	10	BE138336
37	549	12.8	574	9	AI116021
38	548	12.8	432	10	AM258861
39	547	12.7	865	13	BI825221
40	544	12.7	965	13	BI453005
41	543.5	12.7	775	13	BI103238
42	542	12.6	628	13	BM487251
43	541.5	12.6	724	13	BM290186
44	539	12.6	418	10	BB687602
45	538	12.5	749	10	AM026258

## ALIGNMENTS

RESULT 1  
LOCUS BC032938 3337 bp mRNA linear HTC 17-JUN-2002  
DEFINITION Homo sapiens, angiotensin I converting enzyme (peptidyl-dipeptidase  
ACCESSION BC032938  
VERSION BC032938.1 GI:21432080  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 3337)  
Strausberg, R.  
TITLE Direct Submission



## JOURNAL

Submitted (14-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs@email.nih.gov](mailto:cgabbs@email.nih.gov)  
Tissue Procurement: Miklos Palakovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
<http://www.systemsbio.org>  
contact: [amadansystemsbiology.org](mailto:amadansystemsbiology.org)  
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whitting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 34 Row: 1 Column: 13  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 11225608  
This clone has the following problem: frame shifted.

## FEATURES

## SOURCE

1. 3337  
/organism="Homo sapiens"  
/db\_xref="LocusID:59272"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4830668"  
/tissue\_type="Testis"  
/clone\_id="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"

BASE COUNT 1023 a 634 c 756 g 924 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 3337  
Score: 4249.00 Matches: 802  
Percent Similarity: 99.75% Conservative: 1  
Best Local Similarity: 99.63% Mismatches: 2  
Query Match: 99.02% Indels: 1  
DB: 11 Gaps: 0

US-09-978-385-2 (1-805) x BC032938 (1-3337)

Qy 1 MetSerSerSerSerTripleuLeuSerleuValAlaValThAlaAlaGlnSerThr 20  
Db 40 ATGTCAAGCTCTCTGCTGCTCTCTCAAGCTTGTGTGTAACAGTGTCTGCTCAACC 99  
Qy 21 TleGlUGlUGlAlaLysThrPheLeuAspLysPheAsnHisGluAlaGlnAspLeuPhe 40  
Db 100 ATTGAGGAACAGGCAAGCAATTTGGACAAGTTAACCAAGCAAGCAAGCAAGCTGTTTC 159  
Qy 41 TYRGlSerSerleuAlaSerTriPAsnTYRAsnThrAsnIleThrGluGluAsnValGln 60  
Db 160 TATCAAAAGTCACTTCTCTGGAATTTAACAACAATATTACGAAGAAGATGCCAA 219  
Qy 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrleuAla 80  
Db 220 AACATGAATTAATGCTGGGCAAAATGTGCTGCTTTTAAAGAACAGTCCACACTTGCC 279  
Qy 81 GlnMetTYRProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100  
Db 280 CAATGTATCCACTACAAAGAAATTCAGAAATTCACAGTCAAGCTTCAGTGGAGGCTTT 339  
Qy 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120  
Db 340 CAGCAAAATGCGTCTCAAGTCTCTCAGAAAGCAAGCAAGCAAGCGTTGAACAACATTTCA 399  
Qy 121 AsnThrMetSerThrIleTYRSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140  
Db 400 AATACAAATGAGCACCATCTACAGTACGAAAGATTGTAAACCAAGATATCCACAAGAA 459

Qy 141 CysLeuLeuLeuGlnProGlyLeuAsnGluIleMetAlaAsnSerleuAspTYRAsnGlu 160  
Db 460 TGCCTATTACTTACACAGAGTTGAATGAATGAAGCAAGTTTATACATCAATGAG 519  
Qy 161 ArgLeuTriPAlaTriPLeuSerTriPArgSerGluValGlyLeuGlnLeuArgProLeuTYR 180  
Db 520 AGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGGTGGCAAGCAGCTGAGCCATATAT 579  
Qy 181 GluGluTYRValValLeuLysAsnGluMetAlaAlaGlnAsnHisTYRLeuAspTYRGLY 200  
Db 580 GAAGAGATGTGCTCTGAAATAATGAGATGGCAAGGCAAAATCATTTATGAGACATAGGG 639  
Qy 201 AspTYRTriPArgGlyAspTYRGLYValAsnGlyValAspGLYTYRAspTYRArgGly 220  
Db 640 GATTATTGGAGAGAGAGCTATGAAATGAATGGGTGATGATGCTATACATACAGCCGGCC 699  
Qy 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTYRGLYHisLeu 240  
Db 700 CAGTGTATTGAAGATGGGAAACATACCTTTGAAGAGATTAACCATTAATGACAATCTT 759  
Qy 241 HisAlaTYRValAlaGlnAlaLysLeuMetAsnAlaTYRProSerTYRTrIleSerProIleGly 260  
Db 760 CATGCCATGTGAGGCGCAAGTTGATGAATGCTATCTCTTAATCAGATCCAAATTGGA 819  
Qy 261 CysLeuProAlaHisLeuLeuGlyAspMetTriPArgPheTriPAsnLeuTYRSer 280  
Db 820 TGCCTCCCTGCTCATTTGCTTGGTGAATGTGGGGATGAGATTTTGGACAATGCTACTCT 879  
Qy 281 LeuThrValProPheGlyGlnLysProAsnIleAspValIthrAspAlaMetValAspGln 300  
Db 880 TTGCACATTCCTCTTGGACAGAAACCAACATGATGTTACTGATCAATGGTGGGACAG 939  
Qy 301 AlaTriPAspAlaGlnArgGluIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320  
Db 940 GCGTGGATGACAGAGATATA-TTCAGAGAGCGGAGAGTCTTTGTATCTGTGGTCTT 998  
Qy 321 ProAsnMetThrGlnGlyPheTriPGLYAsnSerMetLeuThrAspProGlyAsnValGln 340  
Db 999 CCTAATATGACTCAAGATTCCTGGGAAATTTCCATGCTAACGAGCAAGCAAGATGTTCA 1058  
Qy 341 LysAlaValCysHisProThrAlaTriPAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
Db 1059 AAACAGCTGTGCTATCCACAGCTTGGGACCTGGGAGGCGGCACTTCAGATCTTATG 1118  
Qy 361 CysThrLysValIthrMetAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380  
Db 1119 TGCACAAAGGTGACAAATGGAGAGCTTCTGACAGCTATCATGAGATGGGCAATATCCAG 1178  
Qy 381 TYRAspMetAlaTYRAlaAlaGlnProPheLeuLysAsnGlyAlaAsnGluGlyPhe 400  
Db 1179 TATGATATGGCATATGCTGCACAACTTTTCTGTAAGAAATGAGCTAATGAAGATTC 1238  
Qy 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaIleThrProLysHisLeuLysSer 420  
Db 1239 CATGAAGCTGTGGGGAATCATCTTCCTTGCACCCACCTTAAGCAATTTAAATTC 1298  
Qy 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeu 440  
Db 1299 ATTGGTCTTCTGTACCCGATTTTCCAAAGAACATGAACAAATAAACTTCGCTC 1358  
Qy 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTYRMetLeuGlyLysTYRArg 460  
Db 1359 AAACAGACACTCAGATTTGGGAGCTGCTGCATTTACTTAACATGTTAGAGAGTGAGG 1418  
Qy 461 TrpMetValPheLysGlyGluIleProLysAspGlnTriPMetLysLysTriPArgLysMet 480  
Db 1419 TGGATGTGCTTTAAAGGGGAAATTCCAAAGCAAGTGAGTGAAGAAAGTGTGGAGATG 1478  
Qy 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTYRAspPro 500  
Db 1479 AAGCGAGATAGTTGGGGTGTGGAACTGTGCACCATGATGAAACATATCTGACACCC 1538

QY	501	AlaSerLeuPheHisValSerAsnAspTrySerPheIleArgTyrThrArgThrLeu	520
Db	1539	GCATCTCGTCCATCGTTTCTTAATGATTAATCTCATTCATTCGATATTATACAAAGAGCCCTT	1598
QY	521	TyrGlnPheGlnPheGlnGlnLualLeuGlyGlnAlaIleAlaIleHisGlnGlyProLeuHis	540
Db	1599	TACCAATTCACGTTTCCAAAGAACACTTTGTCAAGCAGCGCTAAACATGAAGAGGCCCTCGCAC	1658
QY	541	LysCysAspLlSerAsnSerThrGluAlaGlnGlnLysLeuPheAsnMetLeuArgLeu	560
Db	1659	AAATGTGACATCTCAAACTCTACAGAAAGCTGGACAGAAACAGTTTCAATATCTGAGGCTT	1718
QY	561	GlyLysSerGlnProTrpThrLeuAlaLeuGlnAsnValValGlyAlaLysAsnMetAsn	580
Db	1719	GGAAATTCAGAACCCCTGGAGCCCTGACATTTGGAAAATGTTGTTAGAGACCAAGAACATGAAAT	1778
QY	581	ValArgProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLys	600
Db	1779	GTAAGGCCACATCGCTCAACTACTTTAGAGCCCTTATTTACCTGGCTGGAAGACCAAGAACAG	1838
QY	601	AsnSerPheValJalYrrPserThrAspTrpSerProTyrAlaAspGlnSerIleLysVal	620
Db	1839	AATTCCTTTGTGGAGGAGGTACCCAGACGAGATCCATGTATGCAGACCAACATCAAGAGTG	1898
QY	621	ArgLlSerLeuLysSerSerAlaLeuGlnLysPylsAlaTyrGluTrpAsnAspAsnGlnMet	640
Db	1899	AGGATTAAGCCTAAATCAACACTCTTGTGAAATTAACCATATGAATGGAACGACATGAAATG	1958
QY	641	TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn	660
Db	1959	TACCTCTTCGATCACTGTGTGCATATCTATGAGGACAGTACTTTTAAAGGAAATAAT	2018
QY	661	GlnMetLlLeuPheGlnGlyGlnLysPylsAlaTyrValAlaAsnLeuLysProArgLlSer	680
Db	2019	CAGATATTCCTTTTGTGGAGAGAGATGTGGAGTGGCTAATTTGAAACCAAGAAATCTCC	2078
QY	681	PheAsnPhePheValThrAlaProLysAsnValSerAspLlLeuLeuProArgThrGluVal	700
Db	2079	TTTAATTTCTTTGTCTACGACACATAAAATGTGTCTGATATTCATCTCTGAACCTGAATTT	2138
QY	701	GluLysAlaIleLeuGlnMetSerArgSerArgLlAsnAspAlaPheArgLeuAsnAspAsn	720
Db	2139	GAAAGAGCCATCAGAGATGCCCGGAGCCCTATCAATGAATGATTCGCTCGTCAATGACACAC	2198
QY	721	SerLeuGlnPheLeuGlyLlGlnProThrLeuGlyProProAsnGlnProProValSer	740
Db	2199	AGCCTAGAGATTCTGGGGATACAGCCCAACACTTGGACCTCCCTACACAGCCCTGTTTCC	2258
QY	741	LleTrpLeuLlValAlPheGlyValValMetGlyValLlValAlaGlyLlValAlleu	760
Db	2255	ATATGGCGATTTGTTTTTGGATTTGAGTGGAGGAGTATGATGTGTTGGCATTTGTATCTCTG	2318
QY	761	LlePheThrGlyLlLeuAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro	780
Db	2319	ATCTTCACTGGGATCAGACAGATCGGAAAGAAATAAATAAAGCAAGAGGGGAAATAATCTT	2378
QY	781	TyrAlaSerLlLeuAspLlSerLysGlyLysAsnProGlyLysPheGlnAsnThrAspAsp	800
Db	2379	TATCGCTCCATCGATTATTACCAAGAGAAATAATATCCAGATTCCTCAAAACACTGATGAT	2438
QY	801	ValGlnThrSerPhe	805
Db	2439	GTTCAAGACTCTTT	2453
RESULT 2			
LOCUS	AL551235		
DEFINITION	AL551235 LMI_NFL006.PL2 Homo sapiens cDNA clone CSOD1041Y14.5	987 bp	mRNA linear EST 16-FEB-2001
ACCESSION	AL551235		prime, mRNA sequence.
VERSION	AL551235.1	GI:12889986	
KEYWORDS	EST.		
SOURCE	human.		

QY 688 ProlysaValSerAspIleIleProArgThrGluValGluValAlaIleArgMetSer 707  
 DB 483 CTTAAATAAGTGTGATATCTATCTCTAGAACTGAAAGTGAAGCCATCAGATGACC 542  
 QY 708 ArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLeuGlyLe 727  
 DB 543 CGAGAGCGATATCAATGATCTCTCCGTCTGATGACACAGCTGAGCTTCTGGGGATA 602  
 QY 728 GlnProThrLeuGlyProProAsnGlnProProValSerIleTriePheLeuIleValPheGly 747  
 DB 603 CACCAACACTGACCTCTTACACGACCCCTGTTTCATATGCGTATGTTTGTGA 662  
 QY 748 ValValMetGlyValIleValValGlyIleValIleLeuIlePheThrGlyIleArgasp 767  
 DB 663 GTTGTATGGAGTATGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 722  
 QY 768 ArgLysLysLysLysAsnLysAlaArgSerGlyGluAsnProTyrAlaSerIleAspSer 787  
 DB 723 CGGAAAGAAAGAAATAAAGCAAGAGTGAAGAAATCTTATGCTCCATCATTTTAC 782  
 QY 788 LysGlyGluAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPhe 805  
 DB 783 AAAGGAAATAATATCATCAGATTCACAAACACATGATGATGCTGACACTCTT 836

RESULT 3  
 AK008530  
 LOCUS  
 DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length  
 CONVERTING ENZYME-LIKE PROTEIN (ACE-RELATED CARBOXYPEPTIDASE ACE2),  
 full insert sequence.  
 AK008530  
 AK008530.1 GI:12842766  
 HTG: CAP trapper.  
 Mus musculus (strain: C57BL/6J) adult male small intestine cDNA to  
 mRNA, clone lib: RIKEN full-length enriched mouse cDNA library  
 clone: 2010305L05.

ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 2049374  
 11042159

REFERENCE  
 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Kashiwagi, K.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 20330913  
 11076861

REFERENCE  
 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
 Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,  
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,

Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
 Quackenbush, J., Schiml, L.M., Straub, F., Suzuki, R., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Baldarelli, R., Barish, G., Blake, J., Boileau, J.D., Bojunga, N.,  
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,  
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
 Hofmann, M., Hume, D.A., Kamlay, M., Lee, N.H., Lyons, P.,  
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombærts, P., Nordone, P.,  
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
 Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
 Toyooka, K., Wang, K.H., Weltz, C., Whitaker, C., Wilting, L.,  
 Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S.  
 and Hayashizaki, Y.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 21085660  
 11217851

REFERENCE  
 5 (bases 1 to 1603)  
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
 Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bull, C.,  
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,  
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,  
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,  
 Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,  
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,  
 Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,  
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
 Schiml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
 Tanaka, T., Teijima, Y., Toyama, T., Yamamura, T., Yamana, I.,  
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
 Hayashizaki, Y.  
 Direct Submission  
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216)  
 Please visit our web site (http://genome.gsc.riken.go.jp/) for  
 further details.

COMMENT  
 FEATURES  
 source

CDS  
 1..1603  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM:DB:2010305L05"  
 /db\_xref="MGI:1902242"  
 /db\_xref="taxon:10090"  
 /clone="2010305L05"  
 /sex="male"  
 /tissue\_type="small intestine"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="adult"  
 49..798

/note="data source: SFR, source key: Q9NRA7, evidence: ISS  
 homolog to ANGIOGENIN CONVERTING ENZYME-LIKE PROTEIN  
 (ACE-RELATED CARBOXYPEPTIDASE ACE2)  
 putative"  
 /codon\_start=1  
 /protein\_id="BAB25723.1"  
 /db\_xref="GI:12842766"  
 /db\_xref="MGI:1917258"  
 /translation="MISGENSEPTKALENVGARNVDYPLNYPQPLFEDMKEQNR  
 NSFVGNTPSPVADSDIKVRSISKALGANAAEWENEFLEPSSVAAMARKVEFII  
 KNQVPLFEDYRVSDDLKPRVSPFEPSPONSDVDPRESDAIRMGRINDVFG  
 LINDNLEFLGHPLEPTQPPVTHILITGYSVALVYVGIILLYTGKGRKKMET  
 KRENSYSDMDIGGESNAGFQNSDDQVTS"  
 1584..1589  
 /note="putative"  
 1603

polya\_signal  
 polyA\_site

/note="putative"

BASE COUNT 465 a 321 c 366 g 450 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 1.44e-118 Length: 1603  
Score: 1071.00 Matches: 200  
Percent Similarity: 87.50% Conservative: 31  
Best Local Similarity: 75.76% Mismatches: 33  
Query Match: 24.96% Indels: 0  
Gaps: 0

US-09-978-385-2 (1-805) x AK008530 (1-1603)

QY 542 CysAspIleSerAsnSerThrGluAlaGlyGlnIleuPheAsnMetLeuArgLeuGly 561  
Db 4 TGAGACATCTCAATTCACAGTGGCGACAGAGTGTCTCAAGATGCTGAGCTTGA 63  
QY 562 LysSerGluProThrPheLeuAlaLeuGlnAsnValAlaGlyAlaLysAsnMetAsnVal 581  
Db 64 AATTGAGAGCCCTGAGCAAGCCCTGGAAATGTGTGAGAGCAAGCAATATGATGTA 123  
QY 582 ArgProLeuAsnThrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsn 601  
Db 124 AATCCACTGCTCAATTAATCTCCACCGTGTGTGACTGGCTGAAGACAGACAGAAAT 183  
QY 602 SerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysValArg 621  
Db 184 TCTTTTGTGGGTGCAACACGATGAGAGCCCATATGCCAGCAACCAATTAAGATGAG 243  
QY 622 IleSerLeuLysSerAlaLeuGlnLysPlyAlaTyrGluTrpAsnAspAsnGluMetTyr 641  
Db 244 ATTAGCCCTTAATACAGCTCTTGTGACCTAATGATGATGATGATGATGATGATGATG 303  
QY 642 LeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGln 661  
Db 304 CTGTCCGATCATCTGTGCATATGCCATGAGAAAGTATTTTCAATATATCAAAACACAG 363  
QY 662 MetIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProArgIleSerPhe 681  
Db 364 AAGATTCCTTTCTAGAGGAGAGATGTAGAGTGTGATGATGATGATGATGATGATGATG 423  
QY 682 AsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTrpGlnValGlu 701  
Db 424 TACTCTTGTGACCTCAACCCCAAAATGTCTGATGATGATGATGATGATGATGATGATG 483  
QY 702 LysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSer 721  
Db 484 GATGCCATCAGATGTCTCTGGGGCCGATCAATGATGCTTGGCCCTGGAATGATTAACAGC 543  
QY 722 LeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerIle 741  
Db 544 CTGGAGTTTCTGGGATTCACCCCAACACTGAGCCACCTTCACAGCCCTCTGTCACATA 603  
QY 742 TrpLeuIleValPheGlyValAlaMetGlyValIleValAlaGlyIleValIleLeuIle 761  
Db 604 TGGCGATTAATTTTGGGTGTGTGATGAGCACTGCTGATGCTGATGCTGATGCTGATG 663  
QY 762 PheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnProTyr 781  
Db 664 CTCACGTGGATCAAGGTGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCTTTAT 723  
QY 782 AlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAspVal 801  
Db 724 GATCGCATGAGCATTTGGAAAGGAAAGGAAAGCAATGCAAGATTCAAAACAGTATGCT 783  
QY 802 GlnThrSerPhe 805  
Db 784 CAGACTTCCTTT 795

RESULT 4  
BI913504 763 bp mRNA linear EST 16-OCT-2001  
LOCUS BI913504  
DEFINITION 603179133F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5243048 5',

## ACCESSION

BI913504.1 GI:16177911

## KEYWORDS

EST.

## SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 763)

NIH-MGC <http://mcc.ncl.nih.gov/>.

REFERENCE

1 (bases 1 to 763)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: [cqabbs-f@mail.nih.gov](mailto:cqabbs-f@mail.nih.gov)

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LMNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L16M1613 row: b column: 09

High quality sequence stop: 675.

Location/Qualifiers

1..763

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5243048"

/lab\_host="DH10B"

/note="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH-MGC Library."

BASE COUNT 237 a 167 c 170 g 189 t

ORIGIN

Alignment Scores:

Pred. No.: 4.74e-117 Length: 763  
Score: 1054.50 Matches: 225  
Percent Similarity: 91.30% Conservative: 6  
Best Local Similarity: 88.93% Mismatches: 12  
Query Match: 24.57% Indels: 10  
DB: 13 Gaps: 3

US-09-978-385-2 (1-805) x BI913504 (1-763)

QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaGlnSerThr 20  
Db 19 ATGTCAAGCTCTCTCGGCTCTCTCAGCTTGTCTGTAAGTCTGCTGATGCTCACC 78  
QY 21 Ile-GluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGlnAspLeuPhe 40  
Db 79 ATGTGAGGAAAGGCAAGACATTTTGGACAAGTTTAACACGAGGACGGAAGACTGTT 138  
QY 40 eTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGlnGlnAsnValG 60  
Db 139 CTATCAAGTCTCACTCTCTTGGAAATTATACCAACCAATATATACGAGAAATCTCCA 198  
QY 60 nasMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAl 80  
Db 199 AAACATGAATAATAGCTGGGACAAATGCTGCTTTTAAGGAAGATCCACACTTGC 258  
QY 80 aglMetTyrProLeuGlnGluTyrIleGlnAsnLeuThrValLysLysGlnIleAlaLe 100  
Db 259 CCAATATATTCACATCAAGAAATTCAGATATTCACAGTCAAGCTTCAAGTCAAGCTCT 318

QY 100 ucIngluansnglySerValIleuSerGluAspLysSerLysArgLeuAsnThrIlele 120  
 |||||  
 Db 319 TCAGCAAAATGSGCTTCTAGGCTCTCGAGACAGCAAGCGTTGAACAATCTT 378  
 QY 120 uAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnI 140  
 |||||  
 Db 379 AATATCAATGACACACATCTACAGTACGAAAAGTTTGTAAACCAATATATCCACAGA 438  
 QY 140 ucYrsIleuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnG 160  
 |||||  
 Db 439 ATGCTTATTAATTCGACCAAGGTTTGACCTACATTAATGCAACAGCTTTAGACTCAATGA 498  
 QY 160 uArgLeuThrPalatp---GluSerTrrPargSerGluValGlyLysGlnLeuArgProle 179  
 |||||  
 Db 499 GAGGCTCGGCTGGCCAAACGCTGCAGATCTGAGGTGCGCAAGCAGCTGAGGCGCAT 558  
 QY 179 uYrGluGluTyrValValIleu-LysAsnGluMetAlaArgAlaAsnHisTyrGluSpt 199  
 |||||  
 Db 559 ATATGAAGAGTATGTGTGTAACAAACATGAGATGGCAAGCAATCATTAAGAGACT 618  
 QY 199 YrGlyAspTrrParg-GlyAspTyrGluVal-AsnGlyVal-AspGlyTyrAspTyrS 218  
 |||||  
 Db 619 ATGGGATTAATTCGAGAGAGAGACATTAAGTACATGCGGTAGCAGCTATGACTAC 678  
 QY 218 eArgGly-GlnLeuIleGlu---AspValGluHisThrPhe---GluGluIleLysPro 235  
 |||||  
 Db 679 GCGCGGACCGATTAATTCGACAGCTGTGAGACATACCTTTGACGAGACTTAAACCA 738  
 QY 236 LeuTyr-GluHisLeuHisAlaTyr 243  
 |||||  
 Db 739 TTTACTGACACCTCTTCATGCCAT 763  
 RESULT 5  
 BG401683 879 bp mRNA linear EST 12-MAR-2001  
 LOCUS 60246350F1 NIH\_MGC\_75 Homo sapiens cDNA clone IMAGE:4594140 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG401683  
 VERSION BG401683.1 GI:33295131  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 879)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: rgs@bhs-remail.nih.gov  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L10M1335 row: d column: 13  
 High quality sequence stop: 619.  
 FEATURES  
 source location/Qualifiers  
 1..879  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4594140"  
 /clone.lib="NIH\_MGC\_75"  
 /lab.host="DH10B (TI phage-resistant)"  
 /note="Organ: Kidney; Vector: pDNR-LIB (Clontech); Site\_1:  
 SfiI (ggcgccctggcc); Site\_2: SfiI (ggccattatggcc); 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CACGGCATTAATGCGC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCGGAGCGGCGGACATG-drr(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.65  
 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts

by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH MGC Library."

BASE COUNT 275 a 177 c 210 g 216 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,15e-116 Length: 879  
 Score: 1048.50 Matches: 218  
 Percent Similarity: 80.50% Conservative: 9  
 Best Local Similarity: 77.30% Mismatches: 12  
 Query Match: 24.43% Indels: 45  
 DB: 12 Gaps: 4

US-09-978-385-2 (1-805) x BG401683 (1-879)

QY 388 GlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPheHisGluAlaValGlyGluIle 407  
 |||||  
 Db 2 CAACCTTTCTGCTAGAGAAATGAGACTATGAAAGATTCCATGAACTGTGGGGAATC 61.  
 QY 408 MetSerLeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAsp 427  
 |||||  
 Db 62 ATGTACCTTCTGCAGCCACACCTAAGCATTTAAATCCATGAGTCTTGTCAACCGAT 121  
 QY 428 PheGlnGluAspAsnGluThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIleVal 447  
 |||||  
 Db 122 TTTCAAGAAAGAAATGAAGAAATTAACCTTCTGCAACACAGCTCAGCATTTGT- 180  
 QY 448 GlyThrLeuProPheThrTyrMetLeuGluLysTrrPargTrrPheValPheLysGlyGlu 467  
 |||||  
 Db 181 GGGACTCTGCCATTTACTTACATGTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 QY 468 IleProLysAspGlnTrrPheMetLysLysTrrPheMetLysArgGluIleValGlyVal 487  
 |||||  
 Db 241 ATTCCCAAGACACCTGATGATGAAAGAGTGGGAGAGATGAG- 282  
 QY 488 ValGluProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHisValSer 507  
 |||||  
 Db 282 ----- 282  
 QY 508 AsnAspTyrSerPheIleArgTyrThrArgThrLeuTyrGlnPheGlnGlu 527  
 |||||  
 Db 283 -----ACCTTTACCAATTCAGTTTCAAGAA 309  
 QY 528 AlaLeuGysGlnAlaAlaLysHisGlnGlyProLeuHisLysCysAspLieserAsnSer 547  
 |||||  
 Db 310 GCACCTTGTCAAGACGCTAAACATGAGAGGCGCTGCAATGTGCAATCTCAAACTCT 369  
 QY 548 ThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProTrrPhe 567  
 |||||  
 Db 370 ACAGAGAGCTGGACAGAAACTGTCAATGTGCTGAGGCTGGAAATCAGAACCTTGAGAC 429  
 QY 568 LeuAlaLeuGluAsnValValGlyAlaLysAsnMetLysValArgProLeuLeuAsnTyr 587  
 |||||  
 Db 430 CTTCAGCTTGGAAAGTGTGTGAGAGCAAGCAATGATGTAGAGCCAGCTGCAACTAC 489  
 QY 588 PheGluProLeuPheThrTrrPheLysAspGlnAsnLysAsnSerPheValGlyTrrPhe 607  
 |||||  
 Db 490 TTTGAGCCCTTATTTACTGCGGTGAGAACACCAAGAAATGTTTGTGGGATGAGT 549  
 QY 608 ThrAspTrrPheProTrrPheAlaAsp-GlnSerIleLysValArgLieserLeuLysSerAl 627  
 |||||  
 Db 550 ACCGCTGAGGCTATGACAGAACCAAGATCAAGATGAGGATTAATTCAGT 609  
 QY 627 AlaLeuGlyAspLysAlaTyrGluTrrPheAspAsnGluMetLysTrrPheArgSerSerVal 647  
 |||||  
 Db 610 T---GGAGTAAAGCA-TATGACTGAGC---ACATGAATTTCTCTGTCGATCATCTGT 662  
 QY 647 AlaTyr-AlaMetArgGlnTrrPheLysValLysAsnGlnMetIleLeuPheGlyG 667  
 |||||  
 Db 663 GGCATTTCTATTCAGGAGTACTTTTAAAGTAAACATCG-----ATGATCTTGGGG 716  
 QY 667 Lu 667

```

Db      717 AG 718
RESULT 6
BG962298
LOCUS    BG962298
DEFINITION 602827090F1 NCI_CGAP_CO24 Mus musculus cDNA clone IMAGE:4981606 5',
          mRNA sequence.
ACCESSION BG962298
VERSION   BG962298.1 GI:14349935
KEYWORDS EST.
SOURCE    house mouse.
          Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 766)
AUTHORS  NIH-MGC http://mgi.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LHAM10983 row: d column: 23
          High quality sequence stop: 708.

FEATURES
    source
        location/Qualifiers
            1..766
            /organism="Mus musculus"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /clone="IMAGE:4981606"
            /clone_lib="NCI_CGAP_CO24"
            /lab_host="DH10B ('11 phage-resistant)"/
            /note="Organ: colon; Vector: pCW-SF0R6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.6 kb. Constructed by Life
            Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 233 a 156 c 179 g 198 t
ORIGIN
Alignment Scores:
Pred. No.: 3.43e-114 Length: 766
Score: 1031.00 Matches: 210
Percent Similarity: 88.85% Conservative: 21
Best Local Similarity: 80.77% Mismatches: 23
Query Match: 24.03% Indels: 9
DB: 13 Gaps: 0

US-09-978-385-2 (1-805) x BG962298 (1-766)
QY 416 LysHisLeuYsSerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGln 435
Db 2 AAGCATGTGAATCCATTGCTCTTCGCCATTCGATTTCACAGATAGCGAAGACAGAG 61
QY 436 IleAsnHeuLeuYsGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMet 455
Db 62 ATAACCTCTCTACTGAAACGCGCATTTGACATTTGGGAACACTACCGTTTACTTACAG 121
QY 456 LeuGluYsTrpArgTrpMetValPheYsGlyGluIleProYsAspIntrPheYs 475
Db 122 TTAGAGAGAGTGAGGTGATGCTTCGCGGGAATTCACCAAGACAGAGTGATGAAA 181
QY 476 LysTrpTrpGluMetYsArgGluIleValGlyValGluProValProHisAspGlu 495
Db 182 AAGTGGGAGAGAGGCGGAGATCGTTGGTGTGGAGCGCTCTCCATCATGATGAA 241
QY 496 ThrTyrCysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyr 515

```

```

Db      242 ACATACCTGTGACCCCTGCATCTCTGTTCATGTTTCAATGATTCATTCATTCATAT 301
QY 516 TyrThrArgThrLeuYsGlnPheGlnPheGlnGlnGlnGlnGlnGlnGlnGlnGln 535
Db 302 TACACAGAGACCATTTACCAATTCACAGTTTCAAGAGCTCTTTGTCAAGCAGCTAAGTAT 361
QY 536 GluGlyProLeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe 555
Db 362 AATGCTCTCTCACAAATGTGACATCTCAATTCATCTCAAGCTGGGAGCAAGTCTCTC 421
QY 556 AsnMetLeuArgLeuGlyLysSerGluProTrpPheLeuAlaLeuGluLysAsnValAla 575
Db 422 AAGATGCTGAGCTTGTGGAATTCAGAGCCCTGGACCAAGCTTGGAAACATGTGGAG 481
QY 575 ValAlaYsAsnMetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrTyrPle 595
Db 482 ACCAAGAGATATGATGATGATAAAACCACTGCTCATTAATCTCAACCGTTGTTGACAGCT 541
QY 595 UlysAspGlnAsnYsAsnSerPhe-ValGly-TrpSerThrAspTrpSerProTyrAla 614
Db 542 GAAAGAGCAGAACAGAAATTTCTTCAGTGGCGGGAACACTGAAATGAGACCCATATAGCC 601
QY 615 AspGlnSerIleLeuValArgIle-SerLeuYsSerAlaLeuGlyAsp-LysAlaTyrG 634
Db 602 GACCAAGACATTAAGCTGATGAGATTAAGCTTAAGTACGCTCTTGGACCTACATGATAG 661
QY 634 LysTrpAsnAsp-AsnGluMetYsLeuPheArgSerSerValAlaTyrAlaMetArgGln 653
Db 662 ACTGAGCCACAGACGAATATGCTCTGTCGATCATCTGTA-GCATATGCCATGAGAGAC- 719
QY 654 TyrPheLeuYsValLysAsnGluMetIleLeuPheGlyGluLysP 669
Db 720 GTCATGCTCAATATCAAAACAGACAGCTCTTTTCA-GAGAGAGAT 766

RESULT 7
B1561069
LOCUS    B1561069
DEFINITION 603254275F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5296531 5',
          mRNA sequence.
ACCESSION B1561069
VERSION   B1561069.1 GI:15448383
KEYWORDS EST.
SOURCE    human.
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 676)
          NIH-MGC http://mgi.nci.nih.gov/
          National Institutes of Health, Mammalian Gene Collection (MGC)
          Unpublished (1999)
          Contact: Robert Strausberg, Ph.D.
          Email: cgapbs-remail.nih.gov
          Tissue Procurement: Miklos Palokovits, M.D., Ph.D.
          cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
          Toshiyuki and Piero Carninci (RIKEN)
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LHAM11750 row: f column: 20
          High quality sequence stop: 674.

FEATURES
    source
        location/Qualifiers
            1..676
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5296531"
            /clone_lib="NIH_MGC_97"
            /lab_host="DH10B"
            /note="Organ: testis; Vector: pBluescriptR (modified
            pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcag
            ); Oligo-dt primed using primer 5'-TTTATTTTATTTTATTTT-3',
            size-selected for average insert size 2.2 kb and

```







QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160  
 DB 459 TGCTTATCTACTGAAACCAAGCTTTGATGAATATGCAACAGCTTTAGACTACAAATGAG 518  
 QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyGlnLeuArgProLeuTyr 180  
 DB 519 AGGCTCTGGCTGGGAAAGCTGAGATCTGAGGTGGCAAGCAAGCTGAGCCATTATAT 578  
 QY 181 GluGluTyrValValLeuLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyr 199  
 DB 579 GAAGAGTATGTGCTTGAATAATGAGATGCAAGAGCAATCATTTATGAGACTAT 635  
 RESULT 9  
 B1561359 609 bp mRNA linear EST 05-SEP-2001  
 LOCUS 60325121F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:5297380 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1561359  
 VERSION B1561359.1 GI:15448673  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 609)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM1752 row: 3 column: 05  
 High quality sequence stop: 606.  
 Location/Qualifiers  
 source  
 1. 609  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5297380"  
 /clone\_lib="NIH\_MGC\_97"  
 /lab\_host="DH10B"  
 /note="Organ: testis; Vector: pBluescriptR (modified  
 pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag  
 ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTAA-3',  
 size-selected for average insert size 2.2 kb and  
 normalized to 10^5. This is a primary library enriched  
 for full-length clones and constructed using the  
 cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC library."  
 BASE COUNT 193 a 128 c 137 g 151 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.56e-98 Length: 609  
 Score: 901.00 Matches: 186  
 Percent Similarity: 97.40% Conservative: 1  
 Best Local Similarity: 96.88% Mismatches: 2  
 Query Match: 21.00% Indels: 4  
 DB: 13 Gaps: 0  
 US-09-978-385-2 (1-805) x B1561359 (1-609)  
 QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaGlnSerThr 20  
 DB 40 ATGTCAAGCTCTTCTGAGCTCTTCTGAGCTTATGTGTACTCTGCTCAGTCCACC 99

QY 21 IleGluGluAlaValThrPheLeuAspLysPheAsnHisGluAlaGlnAspLeuPhe 40  
 DB 100 ATTAGCAACAGCCGCAAGACATTTGTGGACAAAGTTTAAACACAGCAAGCCAGCTGTTC 159  
 QY 41 TyrGlnSerSerLeuAsnTrpAsnTyrAsnThrAsnIlePheGlnGlnValGln 60  
 DB 160 TATCAAGTTCACCTGCTTCTTGGAAATTAACACCAATTAATTAAGAAAGATGTCNA 219  
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeu-TyrGluGlnSerThrLeuAl 80  
 DB 220 AACATGATATATCTCTGGGCAAAATGCTGCTTTTACAAAGACATCCACTTCC 279  
 QY 80 aglMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuAlaLe 100  
 DB 280 CCAAAATGATCCACTACAGAAATTCAGAAATTCACAGCAAGCTTCAGCTGAGCTCT 339  
 QY 100 uGlnGln-AsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThr-Ile 119  
 DB 340 TCAGCAACAAATGGCTCTCTGAGTCTCTCAGAAACAAAGACAAAGCTTGAACACGAAAT 399  
 QY 120 LeuAsnThrMetSerThrIleTyrSerThrGlyValCysAsnProAspAsnProGln 139  
 DB 400 CTAATATACATATGACCAACATCTACACTGGAAGTTTGAACCCAGATATATCCACA 459  
 QY 140 GluCysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsn 159  
 DB 460 GAATGCTTATCTACTTACACAGCTTGAATGAATGAATGAATGAATGAATGAATGAAT 519  
 QY 160 GluArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyGlnLeuArgProLeu 179  
 DB 520 GAAAGCTCTGGGCTTGGGAAAGCTGAGATCTGAGGTGGAAG-CAGCTGAGGCCATTA 578  
 QY 180 TyrGluGluTyrValValLeuLysAsnGln 189  
 DB 579 TATGAAGAGTATGTGCTTGAATAATGAG 608  
 RESULT 10  
 BF789159 965 bp mRNA linear EST 12-JAN-2001  
 LOCUS 602105037F1 NCI\_CGAP\_K1d14 Mus musculus cDNA clone IMAGE:4223312  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BF789159  
 VERSION BF789159.1 GI:12094195  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 965)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps-remail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM9811 row: 1 column: 09  
 High quality sequence stop: 676.  
 Location/Qualifiers  
 source  
 1. 965  
 /organism="Mus musculus"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4223312"  
 /clone\_lib="NCI\_CGAP\_K1d14"  
 /lab\_host="DH10B (TI phage-resistant)"  
 /note="Organ: kidney; Vector: PCMV-SPORT6; Site.1: NotI;







```

/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPOR6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
BASE COUNT      178 a      111 c      145 g      122 t
ORIGIN

Alignment Scores:
Pred. No.:      3.59e-80      Length:      556
Score:          749.50      Matches:      142
Percent Similarity: 86.02%      Conservative: 18
Best Local Similarity: 76.34%      Mismatches: 25
Query Match:      17.47%      Indels:      1
DB:              13      Gaps:      1

US-09-978-385-2 (1-805) x BM030353 (1-556)
QY 71 AlAphLeuLysGluInSerThrLeuAlaGlnMetTyrProLeuGlnGluLeuGlnAsn 90
Db 2 GCCTTTATGAGAACAGATCCCGGATGCGCAAACTTACTCTGAGAAATTCAGAT 61
QY 91 LeuThValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerValLeuSerGlu 110
Db 62 CTCACACTCAAGCGTCATTTGAAGGCCCTTCAGCATATGAGCAGCTCAGCGCTCGGCA 121
QY 111 AspLysSerLysArgLeuAsnThrLleLeuAsnThrMetSerThrLleTyrSerThrGly 130
Db 122 GAGAGAGCAACAGATTTGACACAGATTTCAATTAATGAGCAGCATCTACAGTCTGCG 181
QY 131 LysValLysAsnProAspAsnProGlnGluCysLeuLeuGlnProGlyLeuAsnGlu 150
Db 182 AAAGTTTGGACCA---AATACAGAGAGAGTCTTACGACTTGAACAGGTTTAAATGAC 238
QY 151 IleMetAlaAsnSerLeuAspTyrAsnGlnGluTyrPalatirPgluSerTyrArgSer 170
Db 239 ATTAATGGAAGAACAGACAGACATCAATAGAGGCTCTGCGTGGGAGAGGCTGAGGCT 298
QY 171 GluValLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGluMet 190
Db 299 GAAAGTTGGCAACCACTAAGGCCATTGTATGAGAGTATGAGTCTTGAAGAAATGAGATG 358
QY 191 AlaArgAlaAsnHisTyrGlnAspTyrGlnAspTyrTyrPargLysAspTyrGluValAsn 210
Db 359 GCAGAGGCCAACCAATTATGAGAGTACGAGGACATTTGAGAGGAGATTATGAGGTGACT 418
QY 211 GlyValAspGlyTyrAspTyrSerArgLysGlnLeuLleGluAspValGluHisThrPhe 230
Db 419 GGGGAGGGGAGACTATGACTACGCGGTGACCAAGTTGATGAAGATGTGGAAACGACCTTT 478
QY 231 GluGluLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsn 250
Db 479 GCAGAGATTAAACCATTTATGAGCAACACTTCATGTTATGTGAGGGCAAAAGTTATGATCAT 538
QY 251 AlaTyrProSerTyrIle 256
Db 539 ACCTACCTTCCTATATAC 556

RESULT 15
Bg772384      741 bp      mRNA      linear      EST 15-MAY-2001
LOCUS      602722306F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4839289 5',
DEFINITION      mRNA sequence.
ACCESSION      Bg772384
VERSION      Bg772384.1 GI:14083037
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1 (bases 1 to 741)

```

```

AUTHORS
TITLE      NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
CONTACT      Robert Strusberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LRAM10774 row: k column: 02
High quality sequence stop: 581.
FEATURES
Source
1..741
location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:4839289"
    /clone_lib="NIH_MGC_97"
    /lab_host="DH10B"
    /note="Organ: testis; Vector: pBluescript (modified
    pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
    ); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTTTNN-3',
    size-selected for average insert size 2.2 kb and
    normalized to R0T 5. This is a primary library enriched
    for full-length clones and constructed using the
    Cap-trapper method (Carninci, in preparation). Library
    constructed by M. Brownstein (NHGRI/NHGR), National
    Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      212 a      188 c      165 g      176 t
ORIGIN

Alignment Scores:
Pred. No.:      1.13e-77      Length:      741
Score:          730.50      Matches:      165
Percent Similarity: 87.88%      Conservative: 5
Best Local Similarity: 85.35%      Mismatches: 11
Query Match:      17.02%      Indels:      13
DB:              12      Gaps:      4

US-09-978-385-2 (1-805) x Bg772384 (1-741)
QY 1 MetSerSerSerSerTyrPleuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
Db 58 ATGTCAAGCTCTCTCGGCTCTTCAGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 117
QY 21 IleGluGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGlnLysPheLeuPhe 40
Db 118 ATTGAGAACAGGCCAACAGACATTTTGGACAACTTTAACCGACGACCGAAGACCTGTC 177
QY 41 TyrGlnSerSerLeuAlaSerTyrPasnTyrAsnThrAsnLleThrGlnGluAsnValGln 60
Db 178 TATCAAAAGTTCACCTGCTTCTTGAATTAACCAACCAATATTCGAGAGATGTCGCA 237
QY 61 AsnMetAsnAsnAlaGlyAspLysTyrPserAlaPheLeuLysGlnGlnSerThrLeuAla 80
Db 238 AACATGATATATCTGTGGGACAAATGCTGCTCTTATACCAAGACAGACACACACTGCGC 297
QY 81 Gln-MetTyrProLeuGlnGlnLleGlnAsnLeuThrValLysLeuGlnAlaLeu 100
Db 298 CAAATATGATATCCACTACAAAGAAATTCAGATTCACAGTCGACGCTGCTGCTGCT 357
QY 100 uGlnGlnAsn---GlySerSerValLeuSerGluAspLysSerLys-ArgLeuAsnThrI 119
Db 358 TCAGCACAACACTGGGCTCTTCAGTCTTCAGAGACAAAGACCAACACGCTTGAACACAA 417
QY 119 leuLeu-AsnThr-MetSerThrLleTyrSerThrGly-LysValLysAsnProAspAsn 138
Db 418 TTCTCAATATACACATAGACACATCTACAGTCTGAGACCACTGTGTATACCGATTAATC 477
QY 138 roGlnGluCysLeuLeuGlnProGlyLeuAsnGlu-IleMetAlaAsnSerLeuAsp 157

```

Db 478 CACAAATGCTTATTACTTACACAGCTTGAATGACACTACAGCAACAGTTTAC 537  
 |||  
 Qy 158 TyrAsn---GluArgLeuTirPalatTirpJuser-TirParSergLValGlyLysGlnLe 176  
 |||  
 Db 538 TACCAACTGACGAGCTCTGGGCTTGGGAAGCCGTGAGAGACT---GCAGGTCTGAGC 594  
 |||  
 Qy 176 uArg-----ProLeuTyrGlu-GluTyrValValLeu 186  
 |||  
 Db 595 ACCTGAGGCCACTTATATGAGAGATATGTGGCTTT 632  
 |||  
 RESULT 16  
 B1826471 800 bp mRNA linear EST 04-OCT-2001  
 LOCUS B1826471  
 DEFINITION 60307617F1 NIH\_MGC\_119 Homo sapiens cDNA clone IMAGE:5167987 5',  
 mRNA sequence.  
 ACCESSION B1826471  
 VERSION B1826471.1 GI:15938021  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 800)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L14M1417 row: 1 column: 20  
 High quality sequence stop: 797.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..800  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="5167987"  
 /clone\_id="NIH\_MGC\_119"  
 /tissue\_type="medulla"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-Sport6; Site: 1; Note:  
 anonymous male age 27. Library is oligo-dt primed and  
 directionally cloned (EcoRV site is destroyed upon  
 cloning). Average insert size 1.3 kb, insert size range  
 0.9-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 013. Note:  
 this is a NIH-MGC Library."  
 BASE COUNT 169 a 244 c 230 g 157 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,16e-72 Length: 800  
 Score: 687.50 Matches: 128  
 Percent Similarity: 68.30% Conservative: 53  
 Best Local Similarity: 48.30% Mismatches: 81  
 Query Match: 16.02% Gaps: 4  
 DB: 13 Indels: 3  
 US-09-978-385-2 (1-805) x B1826471 (1-800)  
 Qy 230 PheGluGluIleLysProLeuTyrGluHisLeuHisAlaTyrValArgAlaLysLeuMet 249  
 |||  
 Db 9 TTCACAGAGCTGACGACCTTACCTCAACCTGCATGCTAGTGGCGGGCCCTGCAC 68  
 |||  
 Qy 250 AsnAlaTyr---ProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuGly 268

Db 69 CGTCACAGGGGGCCAGCAGCATCACTGAGAGGGCCCATTCCTCCTCAGCTGGGG 128  
 |||  
 Qy 269 AspMetTirpGlyArgPheTirpHisLeuTyrSerLeuThrValProPheGlyGln 288  
 |||  
 Db 129 AACATGGGGGCGAGACCTGTCCACATCTAAGACTGTGTGGTGGCTTCCTCAGCC 188  
 |||  
 Qy 289 ProAsnIleAspValTirPaPalMetValAspGlnAlaTirPaPalMetArgIlePhe 308  
 |||  
 Db 189 CCCTCAGTACACACAGGAGCTATGCTTAACAGAGGCTGAGCGCCAGAGAGATT 248  
 |||  
 Qy 309 LysGluAlaGluLysPhePheValSerValGlyLeuProAsnMetThrInclPheTirp 328  
 |||  
 Db 249 AAGAGGCTGATGATTCTTCACCTCCGTGGGCTGCTGCCCTGCTGAGTCTGG 308  
 |||  
 Qy 329 GluAsnSerMetLeuThrAspProGlyAsnValGlnLysAlaValCysHisProThrAla 348  
 |||  
 Db 309 AACAGTGTGCTGTGAGAGCCAGCCAGCGGGGAGGTGTGTGACAGCGCTGGCC 368  
 |||  
 Qy 349 TirPaSerLeuGlyLysGly---AspPheArgIleLeuMetCysThrLysValThrMetAsp 367  
 |||  
 Db 369 TGGGACTTTCACACAGGAGGACTTCGGATCAACAGAGTGCACACCGCTGAAGTGGAG 428  
 |||  
 Qy 368 AspPheLeuThrAlaHisHisGluMetGlyHisIleGlnTyrAspMetAlaTyrAla 387  
 |||  
 Db 429 GACCTGTGTGGTGGCCACAGAAATGGCCACATCCAGATTTCTATGACAGTCAAGAC 488  
 |||  
 Qy 388 GluProPheLeuLeuArgAsnGlyAlaAsnGluGlyPheHisGluAlaValGlyGlu 407  
 |||  
 Db 489 TTACTGTGGCTTGAAGAGGGGTGCCACCCGGCTTCATAGGCAATGGGAGCTGG 548  
 |||  
 Qy 408 MetSerLeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAsp 427  
 |||  
 Db 549 CTACCCCTTCACAGTGTGTACGCCCAAGCACTCAAGTCTCACTGTGAGAGTGGAG 608  
 |||  
 Qy 428 PheGlnIleAspAsnGlnThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIleVal 447  
 |||  
 Db 609 GTGGGAGGAGC---GAGCATGATCACTACTTCTGTGAAGTGGCCCTTGCAAGATC 665  
 |||  
 Qy 448 GlyThrLeuProPheThrTyrMetLeuGluLysTirPaTirpMetValPheLysGlyGlu 467  
 |||  
 Db 666 GCCTTATCCCTTCACACTCACTGCTGATCACTGGGCTGGAGGATATGAT-GGAAGC 724  
 |||  
 Qy 468 IleProLysAspGlnTirpMetLysLysTirpGluMetLysArgGluIleValGlyVal 487  
 |||  
 Db 725 ATCCACCAAGGAGTATTAACAGAGTGTGAGGCTCAGGCTGAAGTACAGGGCCT 784  
 |||  
 Qy 488 ValGluProValPro 492  
 |||  
 Db 785 CTGCCCCCATGCCC 799  
 |||  
 RESULT 17  
 BM460886 1154 bp mRNA linear EST 05-FEB-2002  
 LOCUS BM460886  
 DEFINITION AGENCOURT 6419622 NCI CGAP OV44 Mus musculus cDNA clone  
 IMAGE:5504231 5', mRNA sequence.  
 ACCESSION BM460886  
 VERSION BM460886.1 GI:1850926  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1154)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Aaron Hueh  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1M12145 row: d column: 24

High quality sequence start: 13

High quality sequence stop: 756.

Location/Qualifiers

1. 1154

## FEATURES

BASE COUNT 320 a 232 c 266 g 334 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 1e-71 Length: 1154

Score: 684.00 Matches: 129

Percent Similarity: 85.39% Conservative: 23

Best Local Similarity: 72.47% Mismatches: 26

Query Match: 15.94% Indels: 0

DB: 13 Gaps: 0

US-09-978-385-2 (1-805) x BM460886 (1-1154)

QY 628 LeuGlyaspIysalatyrgIutPaspaspangluMetyrIeuphearservVal 647

Db 22 CTGGAGACCTATCATGATGATGACCAACAGAAATGTCCTGTCGATCATCTGT 81

QY 648 AlatyrlametarngIutyrPheleuIysValIysAnclImetIleuPhegIyGlu 667

Db 82 GCATATCCATGAGAAAGTATTTTCAATATCAAAAACGACAGTTCCTTCTTAGAG 141

QY 668 GluaspValaryValAlaAsnLeuIysProArgIleSerPheaspnephValIthraIa 687

Db 142 GAAGATGTAGAGATGAGATGATTTGAAACCAAGAGTCTCTTCTTCTTCTTCTTCTTCA 201

QY 688 ProlysasnValseraspIleleProArgIthGluValGluIysAlaIleargMetSer 707

Db 202 CCCCAATATGTCGTGATGTCATTCCTAGAAAGTGAAGTGAAGATGCATCAAGATGTCT 261

QY 708 ArgSerArgIleasnaspIaPheargLeuAsnAspAsnSerIeugIupheleuGlyIle 727

Db 262 CGGGGCCGATCATGATGTCCTTGGCTGAATGATTAACGCTGAGATTTCTGGGAT 321

QY 728 GluProthrleuGlyProProAsnGlnProProValSerIletrPheleuIleValPhegIy 747

Db 322 CACCCAAACATTTGAGCCACTTACCAAGCCCTGTCACCATATGAGCTGATTTTGGT 381

QY 748 ValValMetGlyValIleValValGlyIleValIleuIlePheThrGlyIleArgasp 767

Db 382 GTGTGATGCGCTAGTGTGATGTCATCATCTCTGATTTGTCATGGGATCAAGGT 441

QY 768 ArgIysIysIysasnIysAlaArgSerGlyGluAsnProTyrAlaSerIleaspIleSer 787

Db 442 CGAAGAGAAAGAAATGAACAAAGAGAGAACCCCTTATGACATCGATGCGATTTGA 501

QY 788 LysGlyGluAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPhe 805

Db 502 AAGGAGAAAGCAATGACAGATTCACAAACAGATGATGCTCAGCTTCTT 555

RESULT 18

LOCUS B1415293 900 bp mRNA linear EST 14-AUG-2001

DEFINITION 602987289F1 NCI\_CGAP\_Lu33 Mus musculus cDNA clone IMAGE:5143199 5',

ACCESSION B1415293

VERSION B1415293.1 GI:15176216

KEYWORDS EST

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE NIH-MGC http://mcc.ncl.nih.gov/

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE Unpublished (1999)

JOURNAL Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgrabs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L1M11353 row: a column: 24

High quality sequence start: 9

High quality sequence stop: 873.

Location/Qualifiers

1. 900

ORIGIN

Alignment Scores:

Pred. No.: 2.11e-71 Length: 900

Score: 680.00 Matches: 131

Percent Similarity: 65.82% Conservative: 50

Best Local Similarity: 47.64% Mismatches: 88

Query Match: 15.85% Indels: 7

DB: 13 Gaps: 4

US-09-978-385-2 (1-805) x B1415293 (1-900)

QY 343 ValCysHISProThrAlaTrpAspLeuGlyLysGly---AspPheArgIleLeuMetCys 361

Db 12 GTGTGCCATCCCTTACGCTGGAGCTTCTTACAAAGCAAGAGCTTACGATCAACAGCT 71

QY 362 ThrIysValIthrMetAspPheleuThrAlaHisIshIshIshIshIshIshIshIshIsh 381

Db 72 ACCTGTGAGACATGAGAGAGCTGTGTATGACGACCAAGAAAGGGGACATCATCATGAT 131

QY 382 AspMetAlaTyrAlaAlaGlnProPheleuLeuArgasnIyAlaasnGlyIshPheHis 401

Db 132 TTCATGACAGTACAAAGACTTACCGGTACTTCCGGGAGGGTCCAAACCCCTGTTTCAT 191

QY 402 GluAlaValGlyGluIleMetSerLeuSerAlaIalThrProLysHISLeuLysSerIle 421

Db 192 GAAGCTATTGGAG--ATATATGCTCTCTCAGTGTCTACCCCAAGACATCTATACAGTCTC 249

QY 422 GlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGlnIleAsnPheLeuLeuLys 441





```

Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1429 row: C column: 16
High quality sequence stop: 770.
Location/Qualifiers
1..778

```

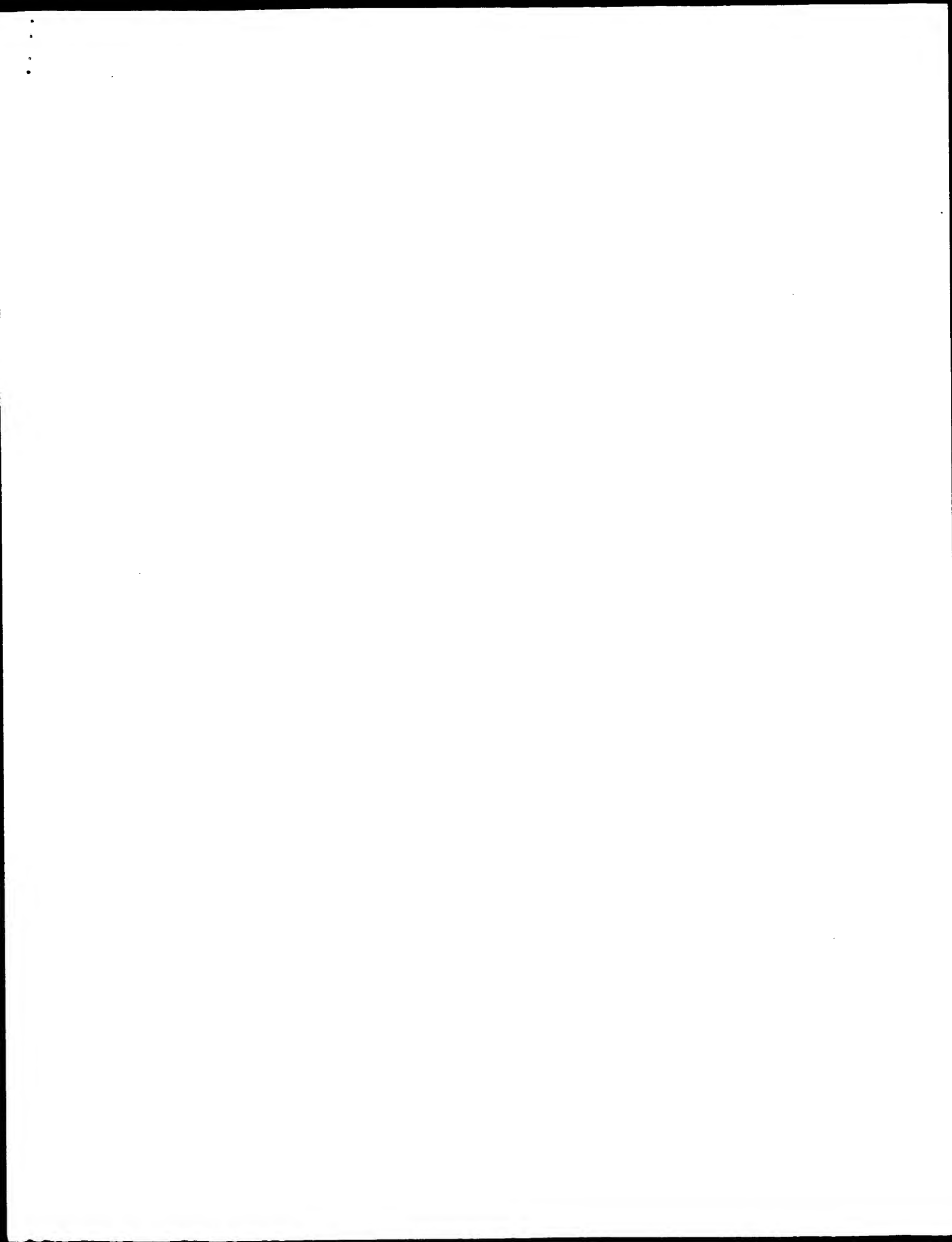
```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5172423"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note this is a NIH_MGC Library."
BASE COUNT      208 a      159 c      161 g      250 t
ORIGIN

Alignment Scores:
  Pired. No.:      1,38e-70
Score:            672.50
Percent Similarity: 97.24%
Best Local Similarity: 95.17%
Query Match:      15.67%
OB:               13

```

Search completed: March 7, 2003, 09:47:36  
Job time : 2390 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 7, 2003, 07:36:48 ; Search time 91 Seconds  
(without alignments)

2712.910 Million cell updates/sec

Title: US-09-978-385-2

Perfect score: 4291  
Sequence: 1 MSSSSWLLSLVAVTAQST.....ISKGNPNPQNTDDVQTSF 805

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODEL=frame.p2n.model -DEV=xlp  
-Q/cgn2.1/USPFO/US09978385/runat.28022003.104717.1368/app-query.fasta.1.967  
-DB-Issued\_Patents.NA -OPMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEXT=0  
-LIST=45 -DOCALIGN=200 -THR.SCORE=pcpt -THR.MAX=100 -THR.MIN=0 -ALIGN=20  
-MODE=LOCAL -OUTPMT=pcpt -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09978385.ecgn.1.1-61-efunat.28022003.104717.1368 -NCPU=6 -ICPU=3  
-NO\_XLPTX -NO\_MAP -LARGEDBDRY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

- Issued\_Patents.NA.\*  
1: /cgn2\_6/ptodata1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata1/ina/5A\_COMB.seq:\*  
4: /cgn2\_6/ptodata1/ina/5B\_COMB.seq:\*  
5: /cgn2\_6/ptodata1/ina/5A\_COMB.seq:\*  
6: /cgn2\_6/ptodata1/ina/5B\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4291	100.0	2415	4	US-08-989-299-3
2	4291	100.0	3396	4	US-08-989-299-1
3	3291	76.7	2350	4	US-09-280-115-40
4	1344	31.3	2478	1	US-08-481-628-1
5	1337	31.2	4020	4	US-09-050-159-130
6	1337	31.2	4024	4	US-09-162-484-18
7	1310	30.5	3942	4	US-09-440-325A-2
8	721	16.8	2082	4	US-08-664-596B-14
9	469	10.9	467	4	US-08-905-223-27
10	381	8.9	848	4	US-09-247-135-27
11	381	8.9	848	4	US-08-157-171-6
12	133.5	3.1	192	1	US-08-157-171-6

13	127	3.0	8700	2	US-08-392-625-16	Sequence 16, Appl
14	127	3.0	8700	2	US-08-466-961A-16	Sequence 16, Appl
15	123	2.9	8700	2	US-08-645-193B-18	Sequence 18, Appl
16	121	2.8	789	4	US-09-280-115-40	Sequence 114, App
17	120.5	2.8	6506	4	US-09-453-702B-11	Sequence 1, Appl
18	120.5	2.8	11384	4	US-08-961-527-45	Sequence 4, Appl
19	120	2.8	144	4	US-08-157-171-4	Sequence 115, App
20	120	2.8	11303	4	US-08-961-527-115	Sequence 4, Appl
21	116.5	2.7	2052	3	US-08-630-916A-45	Sequence 1, Appl
22	116.5	2.7	2052	3	US-08-961-527-115	Sequence 1, Appl
23	113.5	2.6	1998	3	US-08-657-481A-3	Sequence 108, App
24	113.5	2.6	1998	3	US-08-844-059-1	Sequence 4, Appl
25	113	2.6	1998	4	US-09-431-202-1	Sequence 202, App
26	113	2.6	11309	3	US-08-961-527-108	Sequence 3, Appl
27	112	2.6	1839	4	US-09-120-365-4	Sequence 16, Appl
28	112	2.6	1839	4	US-09-515-039-4	Sequence 7, Appl
29	109	2.5	26385	4	US-08-961-527-202	Sequence 33, Appl
30	108.5	2.5	9510	4	US-08-961-527-3	Sequence 3, Appl
31	107	2.5	3425	4	US-09-453-702B-256	Sequence 1, Appl
32	106.5	2.5	37948	4	US-09-107-149-16	Sequence 11, Appl
33	105.5	2.5	3258	4	US-09-251-645-11	Sequence 33, Appl
34	105	2.4	3883	4	US-09-595-424-7	Sequence 3, Appl
35	105	2.4	3883	2	US-08-468-036-33	Sequence 33, Appl
36	105	2.4	3884	4	US-08-376-843-33	Sequence 3, Appl
37	105	2.4	3884	4	US-09-541-782-3	Sequence 3, Appl
38	104	2.4	3672	1	US-09-723-820-3	Sequence 1, Appl
39	104	2.4	3672	3	US-08-491-357-1	Sequence 1, Appl
40	104	2.4	3672	3	US-08-968-633-1	Sequence 1, Appl
41	104	2.4	3672	5	US-09-196-466-1	Sequence 1, Appl
42	101.5	2.4	1683	4	PCT-US96-10823-1	Sequence 1, Appl
43	101.5	2.4	1944	2	US-08-134-001C-316	Sequence 1, Appl
44	101.5	2.4	13977	4	US-08-844-056-1	Sequence 1, Appl
45	101.5	2.4	21706	4	US-09-484-970B-60	Sequence 60, Appl
					US-08-961-527-36	Sequence 36, Appl

## ALIGNMENTS

RESULT 1  
US-08-989-299-3  
Sequence 3, Application US/08989299  
Patent No. 6194556  
GENERAL INFORMATION:  
APPLICANT: Robinson, Susan L.  
TITLE OF INVENTION: ANGIOTENSIN CONVERTING ENZYME HOMOLOG  
NUMBER OF SEQUENCES: 14  
TITLE OF INVENTION: AND THERAPEUTIC AND DIAGNOSTIC USES THEREFOR  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: POLEY, HOAG & ELLIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/989,299  
FILING DATE: 11-DEC-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arnold E., Beth  
REGISTRATION NUMBER: 35,430  
REFERENCE/DOCKET NUMBER: MIA-025.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-7000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2415 base pairs



```

Db 1921 TACCTGTCGATCATCTGTGCATATGCTATGAGGACGACTACCTTTTAAAGTAAAAAT 1980
QY 661 GlnMetIleuPheGlyIleuGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 1981 CAGATGATCTTTTGGGAGGAGATGTGCGAGTGGCTAAATTTGAAACCAAGAAATCTCC 2040
QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
Db 2041 TTTAATTTCTTGTCTACGTCACCTAAATATGTCTCGATATATCTCTTGAACCTGAAGTT 2100
QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2101 GAAAAGCCATCAGATGTCCCGAGCCGTATCAATGATGCTTCGCTGATGACCAAC 2160
QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2161 AGCTTAGAGTTCTGGGGATACAGCCAACTGTGACCTCCCTAACACGCCCCCTTTCC 2220
QY 741 IleThrPheIleValPheGlyValValMetGlyValIleValIleValIleValIleVal 760
Db 2221 ATATGCGCTGATGTTTGTGAGTTGTGATGGAGATGATGTTGCTTCATCTCTG 2280
QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
Db 2281 ATCTTCACCTGGATCAGAGATCGAAGAAAGAAATAAAGCAAGAGTGAGAAATCTCT 2340
QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800
Db 2341 TATGCTCATGATGATTTGCAAGAGAGAAATTAATCCAGATTCACAAACCATGATAT 2400
QY 801 ValGlnThrSerPhe 805
Db 2401 GTTCAGACCTCCTT 2415

RESULT 2
US-08-989-299-1
; Sequence 1, Application US/08989299
; Patent No. 6194556
; GENERAL INFORMATION:
; APPLICANT: Action, Susan L.
; TITLE OF INVENTION: ANGIOGENESIS CONVERTING ENZYME HOMOLOG
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESS: FOLEY, HONG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/989,299
; FILING DATE: 11-DEC-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold E., Beth
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MTA-025.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3396 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 82..2496
US-08-989-299-1

Alignment Scores:
Pred. NO.: 0
Score: 4291.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 4 Gaps: 0

US-09-978-385-2 (1-805) x US-08-989-299-1 (1-3396)
QY 1 MetSerSerSerThrPheLeuLeuSerLeuValAlaThrAlaAlaGlnSerThr 20
Db 82 ATGTCAACCTCTTCTGCTCTCTCTCTCAAGCTTGTCTGTACTGCTGCTCAGTCCACC 141
QY 21 IleGluGluAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
Db 142 ATTAGGAGAACAGGCCAAGACATTTTGGACAAAGTTTAACCAAGAGCCGAAGACTGTTTC 201
QY 41 TyrGlnSerSerLeuAlaSerThrPAsnTyrAsnThrAsnIleThrGluLysAsnValGln 60
Db 202 TATCAAGTTACTTCTGCTCTTGGAAATTAACCAATATTTACTGTAAGAGAAATGTCCAA 261
QY 61 AsnMetAsnAsnAlaGlyAspLysTyrSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 262 AACATGATATATGCTGGGCAAAATGCTGCTCTTTTAAAGAGACAGTCCACACTTGGC 321
QY 81 GlnMetTyrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnIleAlaLeu 100
Db 322 CAATGTATCCACTACCAAGAAATTCAGAAATTCACAGTTCAGCTGACAGCTCTT 381
QY 101 GlnGlnAsnGlySerSerValIleuSerGluAspLysSerLysArgLeuAsnThrIleu 120
Db 382 CAGCAAAATGGGCTTTAGTGTCTGTCTGAGAGACAGCAAGCAAGGTTGAACACAAATTC 441
QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 442 AATACAAATGAGCAGACACTACAGTACAGTGAAGAAAGTTGTAAACCAATATCCACAGAA 501
QY 141 CysLeuLeuLeuGlnProGlyLysAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 502 TGCTTATTTACTTGAAACGAGTTGAATGAATATGCAACAGTTTACACTCAATATGAG 561
QY 161 ArgLeuThrAlaTyrPgiuSerTyrPArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
Db 562 AGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGAGTGGCAAGCAAGCTGAGGCCATTATAT 621
QY 181 GluGluTyrValIleuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
Db 622 GAGAGATATCTGCTTGAAGAAATGAGATGCAAGAGCAAAATCATTAATGAGACTATGG 681
QY 201 AspTyrThrPArgLysAspTyrGluValAsnGlyValAspGlyLysTyrAspTyrSerArg 220
Db 682 GATTATTTGAGAGAGAGACTATGAAGTAAGGGGTGATGCTATGACTATACACCGCGGC 741
QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240
Db 742 CAGTGTATTAAGATGTGGAACATACCTTTGAAGAGATTAACCATTAATATGAACTCTT 801
QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 802 CATGCCATGTGAGGCAAAAGTGAATGCAATGCCATCTCCATATTCAGTCCAAATGGA 861
QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheThrPheHisLeuTyrSer 280
Db 862 TGCTCCCTGCTCATTTGCTGTGATATGTGGGGTAGAATTTTGGCAAAATCTGTACTCT 921
QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300

```

```

Db 922 TTGACAGTTCCTTTGGACAGAAACCAACATAGATGTTACTGATGCAATGGTGGACAG 981
Qy 301 AATATPASPAlaGlnArgIlePheLysGluAlaGluLysPheValSerValGlyLeu 320
Db 982 GCCTGGATGACAGAGATATTTCAAGAGAGCGGAGAAAGTTCTTGATCTGTTGGCTCT 1041
Qy 321 ProAsMetThrGlnGlyPheThrProLysAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 1042 CCTAAATATGACTCAAGAGATTCGGGAAAATTCATGCTTAAACGACCCAGAAATGTTGAG 1101
Qy 341 LysAlaValAlaCysHisProThrAlaThrAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1102 AAGAGAGTGTGCATGCCATCCACACTTGGAGCTGGGGAAAGGGGACTTCAGATCTCTTATG 1161
Qy 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetCysIleGln 380
Db 1162 TGCACAAAGGTGACATGAGACACTTCCTGACAGCTCATGAGATGGGCGATATCCAG 1221
Qy 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
Db 1222 TATGATATGTCATATGCTCACAACTTTCTGCTAAGAAATGAGCTTATAGAGATTC 1281
Qy 401 HisGluAlaValAlaGlyIleLeuMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1282 CATGAGAGTGTGGGAAATCATGTCACCTTCTGCAGCCACACCTAAGCATTTAAATCC 1341
Qy 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440
Db 1342 ATTTGGCTCTTCTGTCACCCATTTTCAAGAGACAAATGAAAGAAATTAACCTCTGCTGC 1401
Qy 441 LysGluAlaLeuThrIleValGlyThrLeuProPheThrMetLeuGluIleThrParG 460
Db 1402 AAACAAGCACTCACGATGTTGGAGCTGCTGCATTTACTTACATGTTAAGAAAGTGGAG 1461
Qy 461 TrpMetValPheLysGlyIleProLysAspGlnTrpMetLysLysTrpIleGluMet 480
Db 1462 TGGATGCTCTTTAAAGGGGAAATTCCTCAAGAGACACAGTGTGATAAAAGTGGGAGATG 1521
Qy 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTrpCysAspPro 500
Db 1522 AACCGAGAGATAGTGGGTGGTGGAGACTGTGCCCATGATGAAACATACACGTGTGACCCC 1581
Qy 501 AlaSerLeuPheHisValSerAsnAspTrpSerPheIleArgTyrTrpThrArgThrLeu 520
Db 1582 GCATCTCTGTCATCTTTCTATATGATTTACTCATTCATTCGATATATACCAAGAGACCTT 1641
Qy 521 TyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaAlaLysHisGluGlyProLeuHis 540
Db 1642 TACCAATTCACAGTTTCAAGAGACACTTTGTCAAGGACAAACATGAAAGCCCTGTGCAC 1701
Qy 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560
Db 1702 AAATGTGACATCTCAAACTCTACAGAGAGCTGGACAGAAACGTGTCAATATGCTGAGGCTT 1761
Qy 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAlaAsnValValGlyAlaLysAsnMetAsn 580
Db 1762 GGAATAATCAGAACCTCGAGACCTAGCATTTGGAATAATGTTGTAGAGCAAAAGACATGAT 1821
Qy 581 ValArgProLeuLeuAsnTrpPheGluProLeuPheThrTrpLeuLysAspGlnLys 600
Db 1822 GTAAGGCACTGCTCAACTTCACTTGGAGCCCTTATTTACCTGGGTGAAACCAAGAACAG 1881
Qy 601 AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
Db 1882 AATCTTTTGTGGATGGAGTACCGACTGAGTCCATATGACAGAACCAAGACATCAAAAGTG 1941
Qy 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640
Db 1942 AGGATTAAGCCTAAATACAGCTTGTGAGATTAAGCATATGAAATGGAACACATGAAATG 2001
Qy 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660

```

```

Db 2002 TACCTGTTCCGATCATCTGTTGCATATGCTATGAGGACGACTTTTAAAGTAAATAAT 2061
Qy 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 2062 CAGATGATCTCTTTTGGGAGGAGATGTCGAGTGGCTAATTTGAAACCAAGAACTTCC 2121
Qy 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleLeuProArgThrGluVal 700
Db 2122 TTTAATTTCTTTGTCATGTCACCTAAATAATCTGTGATATTCATTTCTTACAACTGAAGT 2181
Qy 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2182 GAAAGGCCATCAGAGATGTCCTGGAGCCGTATCATATGATCTTCCGTGAAATGACAAAC 2241
Qy 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2242 AGCTTAAGTCTTGGGGATACACCAACACTTGGACCTCTTAACCGCCCTGTTTCC 2301
Qy 741 IleThrPheIleValPheGlyValAlaMetGlyValIleValAlaGlyIleValIleLeu 760
Db 2302 ATATGGCTGATGTTGTTTGGAGTGTGATGGAGATGATAGTGGTGGCATTTGCATCCTG 2361
Qy 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro 780
Db 2362 ATCTTCACTGGGATCAGAGATCGGAGAGAGAAATTAACCAAGAGTGGAGAAATCCT 2421
Qy 781 TyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAsp 800
Db 2422 TATGCTCTCATCGATATTTAGCAAGAGAAATATATCAGAGATTTCCAAACACTGATGAT 2481
Qy 801 ValGlnThrSerPhe 805
Db 2482 GTTCAGACCTCTT 2496

RESULT 3
US-09-280-116-40/c
: Sequence 40, Application US/09280116A
: Patent No. 6331427
: GENERAL INFORMATION:
: APPLICANT: Robison, Keith E.
: TITLE OR INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
: FILE REFERENCE: 5800-24, 035800/176965
: CURRENT APPLICATION NUMBER: US/09/280.116A
: NUMBER OF SEQ ID NOS: 268
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 40
: LENGTH: 2350
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: angiotensin-converting enzyme
US-09-280-116-40

Alignment Scores:
Pred. No.: 0 Length: 2350
Score: 3291.00 Matches: 726
Percent Similarity: 92.62% Conservative: 2
Best Local Similarity: 92.37% Mismatches: 12
Query Match: 76.70% Indels: 46
DB: 4 Gaps: 2

US-09-978-385-2 (1-805) x US-09-280-116-40 (1-2350)
Qy 1 MetSerSerSerSerTrpLeuLeuLeuSerLeuValAlaVal-ThrAlaAlaGlnSerTh 20
Db 2305 ATGTCAACCTCTCTCTGCTCTCTCTCTCAAGCTTTGTGTGTAGAGCTCTGCTGATGCAC 2246
Qy 20 rIleGluGluGlnAlaLysThr-PheLeuAspLysPheAsnHisGlu-AlaGluAspLeu 39
Db 2245 CATTAGGAGACAGCCAGACATTTTGGACAGATTAAACAGGAAAGCGAAGACCTG 2186
Qy 40 PheTyrGlnSerSerLeuAlaSerTrpAsnTrpAsnThr-AsnIleThrGluGluAsnVa 59

```



Db 2185 TTCATCAAGTCACTGCTTGGAAATTAACCCCAATTTACTGAAAGATGT 2126

Qy 59 IGILSNMETSLSMSALAGLYASP-LysTrpSer-AlaPheLeuLysGluGlnSerThr 78

Db 2125 CCANAAACATGAAATATGCTGGGACAAATGGTGCTGCTTTTAAAGACAGTCCACA 2066

Qy 79 LENAAGLMECTYrProLeuGlnGluLeGILSNMLeuThrValLysLeuGlnLeuGln 98

Db 2065 CTGGCCCAATATGATCCACTACACAGAAATTCAGAAATTCACACTCAAGCTTCATCGAG 2006

Qy 99 ALALEUNGILSNGLYSERSErValLeuSerGluASPysSerLysArgLeuSNrThr 118

Db 2005 GCTCTTAACCAAAATGGTCTTCAGTGCCTCACAAGACAAAGCAACGTTGAAACACA 1946

Qy 119 ILEUASNTHMETSERTHRILE-TyrSerThrGly-LysValCysAsnProASPASP 138

Db 1945 ATCTTAATATCATGTAGCCACCACATCAAGACTGAGAAAAGTTGTAAACCATATATC 1886

Qy 138 ROGLNGIUCYSLSEULEUNGILUPROGLYLEUASNGLILEMETLALSNSERLEUASPT 158

Db 1885 CACAAAGATCTTATTACTTGACACAGGTTGAAATGAATATATGCGAACAGTTTAAGT 1826

Qy 158 IYASNGIUCYSLSEULEUTPALATRPGLISERTPARSERGLUVALGLYSLGILEUARGP 178

Db 1825 ACATGAGAGGCTCTGGCGTTGGGAAAAGCTGGAGAACTCAGCTCGGCAAGCAGCTGAGGC 1766

Qy 178 ROLEUTRGLUQUITYrValValLeuLysASNGLILEMETLARGLALSNSISTRGLUA 198

Db 1765 CATTATATATAAGATATGCTGTGAAAATATAGATGCCAAGACCAAAATATTTATGAG 1706

Qy 198 SPYRTIYASPYRTYrPARGLYASPSPYRGLU-ValASNGLYVAL-ASPGLIYrASPY 217

Db 1705 ACATGGGATTTATGGAGAGAGACATCAAGAAATAGGAGGATTAAGCTATGACATCA 1646

Qy 217 r-SerArgLYSLNLEUILEGLUASPVALGILNHSrThr-PheGluGluILEYSPROLEU 236

Db 1645 CATGCCGCGCCAGCTGATGAAGATGTGGACACATACCTTTTGAGAGATTTAAACCATTA 1586

Qy 237 TYRGLIUNHLSLEUHSIALIATYrVALAGALYSLSEUMETASNALATYrPROSErTYRILE 256

Db 1585 TATGACATCTTCATAGCTTATGTGAGGCGCAAAATGATGATGCTTCCTCTATATC 1526

Qy 257 SERPROILEGLCYSLSEUPROALNHSILEU-LEUGLYSPMETRIPGLY-ArgPheTrp- 275

Db 1525 AQTCCAAATTTGATGCCCTCCCTGCTCATTTTCGTTGGTGAATATGGGGTTAAGATTTTGGG 1466

Qy 276 THRASNLEUYrSERLEUTHrVAL-PROPhEGLYGLN-LYSPROASNILE-ASPVALThr 294

Db 1465 ACAATATCTGATCTTTGACAGTGCCTTCCTTGACAGAGAAACCAACATCATCATTTACT 1406

Qy 295 ASPALA-METVALASPGILNALATRPASALA-GlnARGILEPHELYSGUALA-GLUYS 313

Db 1405 GATGCCAATGTGTGACAGCCTGTGGAGATGCACCAAGAAATTTCAAGGAGGCCAGAG 1346

Qy 314 PHEPHEVALSERVALIYLEUPROASNMErTHrGlnGlyPheTrp-GLASNSErMETILE 333

Db 1345 TTTCTTTATCTCTGCTCTCTCTATATGTAGCTCAAGATTTCTGGGAAAATTCATGCT 1286

Qy 333 UTHR-ASPPROGLYASNVALGILNYSALVALCYSHISPROTHRALATRPASILEU-GLY 352

Db 1285 AAGCGAGCCAGGAATGTTCAGAAAGAGCTGCGCATCCACACTTGGAGACCTGGGGG 1226

Qy 353 LYSGLIYASP-PHEArGILE-LEUMETCYSHIRYrVALrThrMetASPSPheLEUTHrA 372

Db 1225 AAGGGGACATTCAGAGATCTTATATGTGCACAAAGGTGACATGTGACACTTCTGTGACG 1166

Qy 372 LAHSHISGLIUMETGILNHSILIEGLIYrSPMETLArYrALALAGLPROPhELeuT 392

Db 1165 CTCATCATAGATGGGCGCATATCCAGATATGATATGGCATATGCTGCACAACTTTTCTGG 1106

Qy 392 EUARGLASNGLYALASNGLIUCIYPHErISGLIUNLVALGLIUNHMETSERLEUSerA 412

Db	1105	TAAAGAAATGAGACTAAATGAAGATTCCATGAAGCTGTTGGGGAATCATGTCACCTTTCG	1046
QY	412	IAalArthPrC-LysHsLeuLysSerIleGlyLeuLeu---SerProAspPheGlnGlu	430
Db	1045	CAGCCACACCTTAAGACATTAAATATCCATTCGTTGTTTTTTTGTCCACCCGAATTTTTCAGAA	986
QY	431	ASpAsnGluThrPrgLulleAsnPheLeuLeuLysGln-AlaLeuThrIleValGlyThr-L	450
Db	985	GACATGTAAGACAGAAATAAACTTCGCGCAACAAAGACATCAGCATGTGTTGGGACCTCC	926
QY	450	eupProPheThrTYrMetLeuGluLysTPrPAzGlyTrpMetValPheLysGlyGluIleProL	470
Db	925	TGCCATTACTACTACATGTTTAGAAGATGGAGGTGATGGCTTTTAAAGGGGAAATTTCCCA	866
QY	470	ysaAspGlnTrpMetLys-LysTrpTrpGluMetLysArgGluIleValGlyValI---	488
Db	865	AAGACCAAGTGAATGAACAAAGGTGGGAGCATGAAGGAGAGATATGTTGGGTGGTTGGG	806
QY	489	GluProVal-ProHisAspGluThrTYrCysAspProAlaSerLeuPheHisValSerAs	508
Db	805	AACTTGTCGCCCATGATGAACACATCTGTGACCCCGCATCTGTTGTTCCATGTTTCTAA	746
QY	508	naSPtYrSer-PheIleArgTYrThrArgThrLeuTYr-GlnPhe-GlnPheGlnG	527
Db	745	TGATTACTCCATTCATTCCATGATATACAGAAAGACCTTTTAAACCATTCCTCCAGTTCCAGA	686
QY	527	uAlaLeuCys-GlnAlaIleLysHsGlnGlyProLeuHisLysCysAspIleSer-Asn	546
Db	685	AGACCTTTTGCCACACCACTTAACATGAAGGCCCTTGCCAAAGTGAACATCTCGTAAC	626
QY	547	SeTrpThGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProTrp	566
Db	625	TCATACGAAGCGTGAGACAGAACTGTCCATATGCTGAGCTTGGAAATTCAGAACCTGTG	566
QY	567	ThLeuAlaLeuGlnAsnValValGlyAlaLys-AsnMetAsnValArgProLeuLeuAs	586
Db	565	ACCTTGACATTTGGAAATGTTGTAGGACCAAGAACATGATGAAGGCCACATGTCACAA	506
QY	586	n-TyrPheGln-ProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGly	605
Db	505	CGTACTTGTGAGCCGCTTATTACTGTGGCGAAAGACCAAGAAATCTTTTGTGGGA	446
QY	606	TrpSerThrAspTrpSerProTYrAlaAspGln-SerIleLysValArgIle-SerLeuL	625
Db	445	TGGAGTACCGACTGGAGTCCATATGCAGACCAAAAGCATCAAGTGAAGATAAGCTTAA	386
QY	625	ysSerAlaLeuGlnAspLysAlaTYrGluTrpAsnAspAsnGluMetTYrLeuPheArgS	645
Db	385	AATTCAGCTCTTGAGGATTAAGCATATGAATGGAACGAAATAAGTAATGTACCTGTCGAT	326
QY	645	eSerValAlaTYrAlaMetArgGlnTYrPheLeuLysValLysAsnGln-MetIleLeu	664
Db	325	CATCTGTTGCATATGCTATGAGGACAGTACTTTTAAAGATTAATAATTCAGCATGATCTT	266
QY	665	PheGlyGlnGlu-AspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePh	684
Db	265	TTTGGGAGAGAAAGATGTGCGAGTGGCTAAATTTGAACCAGAGATCTCCTTAATTTCTT	206
QY	684	e-ValThrAlaProLysAsnValSerAspIleIleProArgTrpGluValGluLysAlaI	704
Db	205	TGTGATCTGCGCTTAATAATGTGTGTATATCATCTCTGAACATGAAGTTGAAGGGCCA	146
QY	704	leaArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluP	724
Db	145	TCAAGATGTCCGGAGGCGGTATCATATGCTTCCGTGTGAATGAACAAGCCTACAGT	86
QY	724	heLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer	740
Db	85	TTCTGGGGATACAGCCAACACTTGGACCTCTTAACCAAGCCCTGTATTCC	36
RESULT 4			
US-08-481-626-1			
Sequence 1, Application US/08481626			

Patent No. 5801040  
 GENERAL INFORMATION:  
 APPLICANT: Soubrier, Florent  
 APPLICANT: Albert-Gelas, Francois  
 APPLICANT: Hubert, Christine  
 APPLICANT: Corvol, Pierre  
 TITLE OF INVENTION: Nucleic Acid Coding for the Human  
 TITLE OF INVENTION: Testicular Angiotensin Converting Enzyme (ACE) and its  
 TITLE OF INVENTION: Uses, Especially for the In Vitro Screening for this  
 TITLE OF INVENTION: Enzyme in the Organism  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Finegan, Henderson, Farabow, Garrett &  
 ADDRESSEE: Dinner  
 STREET: 1300 I Street, N.W., Suite 700  
 CITY: Washington  
 STATE: D.C.  
 ZIP: 20005-3315  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/481,626  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/656,183  
 FILING DATE: 04-MAR-1991  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: FR 89-09062  
 FILING DATE: 05-JUL-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meyers, Kenneth J.  
 REGISTRATION NUMBER: 25,146  
 REFERENCE/DOCKET NUMBER: 04958-0006-02000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-408-4000  
 TELEFAX: 202-408-4400  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2478 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-481-626-1  
 Alignment Scores:  
 Pred. No.: 2,29e-148 Length: 2478  
 Score: 1344.00 Matches: 259  
 Percent Similarity: 60.97% Conservative: 119  
 Best Local Similarity: 41.77% Mismatches: 204  
 Query Match: 31.32% Indels: 38  
 Gaps: 10  
 US-09-978-385-2 (1-805) x US-08-481-626-1 (1-2478)  
 QY 15 ThrAlaAlaGlnSer-----ThrIleGluGluGlnAlaIleValThrPheLeuAsp 30  
 Db 209 ACATCATGCCCAAGCCCAACCTGCTGACTGATGAGGCTGAGCCACCAAGTTGTGGAG 268  
 QY 31 LysPheAsnHisGluAlaGluAspLeuPheTyrglnSerSerLeuAlaSerTrpAsnTyr 50  
 Db 269 GAATATGACCCGACATCCCGAGGTGTGTGGAACGATGTGCCGAGGCCCAACTGGACTAC 328  
 QY 51 AsnThrAsnIleThrGluGlu-----AsnValGlnAsnMet 62  
 Db 329 AACACCAACATCCACACAGACAGACAGATTCTGTGACAGAGAACATGCAATATAGCC 388  
 QY 63 AsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAlaGlnMet 82

Db 389 AACCAACAC-----CTGAACTAGCGCCAGCCAGAG 424  
 QY 83 TyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeuGln 102  
 Db 425 TTTCATGTGAACAGCTGCAGAACACCAGCTACATCAACGGATCATAAAGAGGTTCAGAC 484  
 QY 103 AsnLysSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeuAsnThr 122  
 Db 485 CTAGAACGGCGAGCGCTGCCCGCCAGAGAGCTGAGAGATACACAAAGATCTGTGGAT 544  
 QY 123 MetSerThrIleTyrSerThrGlnLysValCysAsnProAspAsnProGlnLeuLys 142  
 Db 545 ATGGAACACCACTACAGAGGTGGCCAGCTGTGCCACCCGAATGCC-----AGCTGCTG 598  
 QY 143 LeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspLysArgLeu 162  
 Db 599 CAGCTCGAGCAGATCTGACGAATGTATGCGCCATCCCGGAATATGAAAGCCGTGA 658  
 QY 163 TrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyrGlu 182  
 Db 659 TGGCATGGGAGGCTGGCGAGACAGCGGGGAGAGCCATCCCTCCAGTTTACCCGAGA 718  
 QY 183 TyrValLeuLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGlyAsp 202  
 Db 719 TACGTGACATCATCAACACAGCTGCCGCTCATGAGCTATGATGATGACAGGAGCTCG 778  
 QY 203 TrpArgGlyAspTyrGluValAsnGlyValAspLysTyrAspTyrSerArgGlyGln 222  
 Db 779 TGGAGGTCTATGATACAGACACCATCCCTGGAG----- 811  
 QY 223 IleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGlnHisLeuAla 242  
 Db 812 ---CAAGACCTGAGGGGCTCTCCAGAGCTGCGACCACTCATCTCAACCTGCAAGCC 868  
 QY 243 TyrValArgAlaLysLeuMetAsnAlaTyr---ProSerTyrIleSerProIleGly 261  
 Db 869 TAGGTGGCGGGCCCTGCACCTGACCTAGCGGGCCAGCAACATCACTGAGAGGGCCC 928  
 QY 262 LeuProAlaHisLeuLeuGlnLysPheMetTrpGlyArgPheTrpThrAsnLeuTyrSer 281  
 Db 929 ATTCCTGCTCACCCTGCTGGGAGACATGTGGGGCGAGCTGCTGCACATCTATGACTTG 988  
 QY 282 ThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 301  
 Db 989 GTGTGCTCTTCCTTCACGCCCCCTGATGAGCACACAGAGCATGCTAAAGCAGCGCC 1048  
 QY 302 TrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 321  
 Db 1049 TGGACGCCACAGAGAGATTTAAGAGCCGTGATGATTTCTCACTCCCTGGGGCTGCTG 1108  
 QY 322 AsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 341  
 Db 1109 CCCGTGCTCTGAGTTCTGGACACAGTCTGCTGAGAGAGCAACAGCGGGCGGAG 1168  
 QY 342 AlaValCysHisProThrAlaTrpAspLeuGlyLysGly---AspPheArgIleLeuMet 360  
 Db 1169 GTGTCTGCGCACCGCTCGGCTGAGACTCTACAAACGCAAGAGCTTCGGATCAAGCAG 1228  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHisIleGln 380  
 Db 1229 TGCACACCGGTGAATCTGGAGAGCTGTGTGGCCACACCAAGATGGGCCCATTCAG 1288  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400  
 Db 1289 TATTTCATGCACTAAAGACTTACCTGTGCTGAGGAGGAGGTGCCAACCCCGGCTTC 1348  
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaIleThrProLysHisLeuLysSer 420  
 Db 1349 CATGAGGCAATGTGGGAGCTGTACCTCTCAGTGTGTACGCCCAAGCACTGACAGT 1408  
 QY 421 IleGlyLeuLeuSerProAspPheGlnLysAsnGluThrGlnIleAsnPheLeu 440

```

Db 1409 CTCACCTGCTGAGACAGTGAAGGTGGCAGGAC--GAGCATGACATCACTTTCGTATG 1465
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrIleuGluIleuLysTrpArg 460
||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1466 AAGATGCGCCTTGACAGATCGCCTTATCCCTTACGCTACCTGTCGATCGAGTGGCGC 1525
QY 461 TrpMetValPheIleuGlyGluIleProLysAspGlnTrpMetLysTrpTrpGluMet 480
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1526 TGGAGGTATTTGATGAGACATCACACAGAGACATATACACAGAGTGTGACACCTC 1585
QY 481 LysArgGluIleValGlyValIleGluProValProHisAspGluThrTyrcysAspPro 500
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1586 AGCGTGAAGTACAGAGCGCTCTGCCCCCGACGCTCCAGCATCAAGTCACTTTGACCCA 1645
QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrpThrArgThrLeu 520
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1646 GGGGCCAAGTTCACATCTCTTACGCTGCTTACATCAAGTCACTTTGACCTTCACTC 1705
QY 521 TyrGlnPheGlnPheGlnGlnAlaLeuGlyGlnAlaAlaLysHisGluIleProLysHis 540
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1706 ATCCAGTTCAGTTCACAGAGCATGTGCGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1765
QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1766 AAGGTGACATCTACACAGTCCAGGAGCGCGGCGAGCGCTGGCGAGCCGCAATGAAGCTG 1825
QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValIleGlyAlaLysAsnMetAsn 580
||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1826 GCGTTCAGTGAAGCGCTGCGCGAGAGCCATGCTGATCAAGCGGCAAGCCCAATGATGAC 1885
QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1886 GCGTCGCGCATGTGAGCTACTTCAAGCCGCTCTGAGACGTGCGCGAGGAGAACAG 1945
QY 601 -----AsnSerPheValGlyTrp---SerThrAspTrpSerProTyrAlaAspGlnSer 617
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1946 CTGATGAGGAGAGCTGGCGCTGCGCGAGTACACTGAGCGCGAAGCTCGCTCGCTCA 2005

```

## RESULT 5

```

US-09-050-159-130
; Sequence 130, Application US/09050159A
; Patent No. 6197505
; GENERAL INFORMATION:
; APPLICANT: No. 6197505berg, Ielf T
; APPLICANT: Andersson, Maria K
; APPLICANT: Instrum, Per H
; TITLE OF INVENTION: METHODS FOR ASSESSING CARDIOVASCULAR STATUS AND
; FILE REFERENCE: 1248/ID042
; CURRENT APPLICATION NUMBER: US/09/050,159A
; CURRENT FILING DATE: 1998-03-27
; EARLIER APPLICATION NUMBER: 60/042,930
; EARLIER FILING DATE: 1987-04-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 130
; LENGTH: 4020
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Angiotensin I converting enzyme mRNA
US-09-050-159-130

```

## Alignment Scores:

```

Pred. No.: 3 51e-147 Length: 4020
Score: 1337.00 Matches: 235
Percent Similarity: 61.05% Conservative: 118
Best Local Similarity: 41.73% Mismatches: 204
Query Match: 31.16% Indels: 34
DB: 4 Gaps: 9

```

US-09-978-385-2 (1-805) x US-09-050-159-130 (1-4020)

```

QY 20 ThrIleGluGlnAlaIleThrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 39
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1952 ACTGATGAGCGCTGAGCGCAGCAAGTTGTGGAGGAATATGACCGACATCCAGGTGTG 2011
QY 40 PheTyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGlu----- 57
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2012 TCGAAGAGTATGCGAGCGCAACTGGAATCTACACACCATCATCCACAGAGACAGC 2071
QY 58 -----AsnValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAla 71
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2072 AAGATTCTGCTCAGACAGAAATGCAATGACCAACACACAC----- 2113
QY 72 PheLeuLysGluGlnSerThrLeuAlaGlnMetTyrProLeuGlnIleGlnAsnLeu 91
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2114 -----CTGAATACAGCAGCAGCAGCAGGAAATTGATGATGACCATGTCAGACACC 2167
QY 92 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAsp 111
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2168 ACTATCAAGCGCATCATTAAGAAAGTTTCAGACACTAGAACGGCGAGCGCTGCTGCCAG 2227
QY 112 LysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLys 131
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2228 GAGCTGAGAGAGTACAAACAGATCTCTGTGGATATGAAACACCTACAGCGGTGCGCACT 2287
QY 132 ValCysAsnProAspAsnProGlnGluCysLeuLeuLeuGluProGlyLeuAsnGluIle 151
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2288 GTGTGCGCCAGCAATGCC-----AGCTGCTCGACGCTCGACCCAGATCTACGCAATGTG 2341
QY 152 MetAlaAsnSerLeuAspTyrAsnGluArgLeuThrAlaTrpGluSerTrpArgSerGlu 171
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2342 ATGGCCACATCCCGGAATATGAAGACCGTTATGGCATGGAGGCGTGGCGAGACAG 2401
QY 172 ValGlyLysGlnLeuArgProLeuTyrGluIleTyrValIleLeuLysAsnGluMetAla 191
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2402 GCGGGAGAGCGCATCTCAGTTTACCCGAAATCTGGAACATCAACACAGCGTGC 2461
QY 192 ArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGly 211
||| ||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2462 CGGCTCAATGGCTATGTAGATGACAGGGAGCTCTGAGAGGTCTATGTAGAACACCATCC 2521
QY 212 ValAspGlyTyrAspTyrSerArgLysIleLeuIleGlnAspValGluHisThrPheGlu 231
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2522 CTGGAG-----CAAGACCTGAGAGCGGCTCTTCCAG 2551
QY 232 GlnIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAla 251
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2552 GAGCTGACAGCCTTACCTCAACCTGCATGCTACAGTGCAGCGCGGCGCTGAC 2611
QY 252 Tyr---ProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuGluAspMet 270
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2612 TACGGGGCCGACACATCAACCTGAGGGGCCCATCTCTGCTCACCTGCTGGGAGACATG 2671
QY 271 TrpGlyArgPheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 290
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2672 TGGGGCAGACCTGTGTCACATCATATGACTTGGTGGTGGCCCTTCCCTTACGCCCTCG 2731
QY 291 IleAspValThrAspAlaMetValAspGlnAlaTrpAspAlaGlnArgIlePheLysGlu 310
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2732 ATGGACACCAAGAGAGCTATGCTAAAGCAGGGCTGAGAGCCGCCAGAGAGGATTTTAAAG 2791
QY 311 AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGlyPheTrpGluAsn 330
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2792 GGTATGATTTCTTACCTCCCTGGGGTCTGCGCCGCTGCTGCTGATTTGGAACAAG 2851
QY 331 SerMetLeuThrAspProGlyAsnValGlnLysAlaValCysHisProThrAlaTrpAsp 350
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2852 TCGATGCTGGAAGAACCAACGAGGGGGAGGTGCTGCTCCACAGCTGCGCTGGGAGAC 2911
QY 351 LeuGlyLysGly---AspPheArgIleLeuMetCysThrLysValIleThrMetAspAspPhe 369
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2912 TTCTACAAAGGAGAGTCTCCGATCAAGCAGTGCACACCTGGAACCTTGGAGGACCTG 2971
QY 370 LeuThrAlaHisHisGluMetGlyHisIleGlnTyrAspMetAlaTyrAlaIleGlnPro 389

```

```

Db 2972 GTGGTGGCCACACGAAATGGGACATCCAGTATTCATGATGATACCAAGACTTACCT 3031
    |||
Qy 390 PheLeuLeuArgAsnGlyAlaAsnGluLeuPheHisGluAlaValGlyGluIleMetSer 409
    |||
Db 3032 GTGGCTTGAGGAGGAGGCCAACCCCGCTTCATGAGGCATGAGGAGCTGCTACCC 3091
    |||
Qy 410 LeuSerAlaAlaThrProLysHisLeuLysSerIleGlyLeuLeuSerProAspPheGln 429
    |||
Db 3092 CTCCTAGTGTACGCCCAAGACCTGCACAGTCTCAACCTGCTAGACAGAGAGGTGGC 3151
    |||
Qy 430 GluAspAsnGluThrGluIleAsnPheLeuLeuLysGluAlaLeuThrIleValGlyThr 449
    |||
Db 3152 AGCGAC---GAGCATGACATCAACTTCTGATGAGAGATGGCCCTTGACACAGATCGCCCTT 3208
    |||
Qy 450 LeuProPheThrLeuMetLeuGluLysTrpPheTrpMetValPheLysGlyGluIlePro 469
    |||
Db 3209 ATCCCTTCAGCTACCTGCTGCATCATGAGGCTGGAGGATATTTGATGATGAGATGACC 3268
    |||
Qy 470 LysAspGluTrpMetLysLysTrpPheLysArgGluIleValGlyValValGlu 489
    |||
Db 3269 AAGGAGACTATACAGAGAGAGTGTGAGAGCTCAGGCTGAGAGTACACAGAGCTGCCCC 3328
    |||
Qy 490 ProValProHisAspGluThrTyrcysAspProAlaSerLeuPheHisValSerAsnAsp 509
    |||
Db 3329 CCAGTGCCCGAGGACTCAAGGTGACTTTGACCCAGGCGCAAGTCCACATTCCTTCTAGC 3388
    |||
Qy 510 TyrSerPheLeuArgTyrThrTrpArgThrLeuGluPheGlnPheGlnAlaLeu 529
    |||
Db 3389 GTGCTTACATGAGTACTTCTGCTCAGCTTCATCATCCAGTTCAGTTCACAGAGGACTG 3448
    |||
Qy 530 CysGlnAlaAlaLysHisGluGlyProLeuHisLysCysAspIleSerAsnSerThrGlu 549
    |||
Db 3449 TCCAGGAGCAGGCTGGCCACAGGCGCCCTGCACAGAGTGCATCTACCACTCAAGAG 3508
    |||
Qy 550 AlaGlyGluLysLeuPheAsnMetLeuArgLeuGlyLysSerGluProTrpThrLeuAla 569
    |||
Db 3509 GCCGGGAGAGCGCTGGCCAGCCGCATGAAAGCTTCAGTGAAGCCGTGGCCGGAACCC 3568
    |||
Qy 570 LeuGluAsnValValAlaLysAsnMetAsnValArgProLeuLeuAsnTyrPheGlu 589
    |||
Db 3569 ATGCAGCTGATACAGGCCACGCCCAACATGACAGCCCTGGCCATCTTACGACTCTCAAG 3628
    |||
Qy 590 ProLeuPheThrTrpLeuLysAspGlnAsnLys-----AsnSerPheValGlyTrp--- 606
    |||
Db 3629 CCGCTGCTGAGTGTGCTCCGACAGGAGAACGAGCTGCATGGGAGAACTGGGCTGGCGC 3688
    |||
Qy 607 SerThrAspTrpSerProTyrTrpAlaAspGlnSer 617
    |||
Db 3689 CAGTACACTGGAGCGCCGAACTCCGCTCGCTCA 3721
    |||

```

RESULT 6  
US-09-162-484-18  
Sequence 18, Application US/09162484  
Patent No. 6248724

GENERAL INFORMATION:  
APPLICANT: Phillips, M. Ian  
APPLICANT: Mohuczy, Dagmara  
TITLE OF INVENTION: ANGIOTENSIN Oligonucleotide COMPOSITIONS TARGETED TO  
FILE REFERENCE: UFLA/087/0FLA087P  
CURRENT APPLICATION NUMBER: US/09/162,484  
CURRENT FILING DATE: 1998-09-25  
EARLIER APPLICATION NUMBER: 60/059,661  
EARLIER FILING DATE: 1997-09-25  
NUMBER OF SEQ ID NOS: 19  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 18  
LENGTH: 4024  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-162-484-18

```

Alignment Scores:
Pred. No.: 3,52e-147 Length: 4024
Score: 1337.00 Matches: 255
Percent Similarity: 61.05% Conservative: 118
Best Local Similarity: 41.73% Mismatches: 204
Query Match: 31.16% Indels: 34
DB: Gaps: 9

US-09-978-385-2 (1-805) x US-09-162-484-18 (1-4024)

Qy 20 ThrIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeu 39
    |||
Db 1952 ACTGATGAGGCTAGGCGCACCAAGTTGTGAGGAATATGACCGGACATCCAGGTGG 2011
    |||
Qy 40 PheTyrGlnSerSerLeuAlaSerTyrPheTyrAsnThrAsnIleThrGluGlu----- 57
    |||
Db 2012 TGGACAGATGATGCCAGGCGCACTGGAATCTACACACCAACATCACCACAGAGACCAC 2071
    |||
Qy 58 -----AsnValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAla 71
    |||
Db 2072 AAGATTCGCTGCAGAGAACATGACAAATGACCAACACACC----- 2113
    |||
Qy 72 PheLeuLysGluGlnSerThrIleAlaGlnMetTyrProLeuGlnGluIleGlnAsnLeu 91
    |||
Db 2114 -----CTGAAGTACGCGCACCCAGGACGAGAGTGTGATGATGACCAAGTTCAGAACACC 2167
    |||
Qy 92 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAsp 111
    |||
Db 2168 ACTATCAAGGAGGATCATAAAGAGTTTCAGAGACTTGAACGGGACGCTGCTGCCAG 2227
    |||
Qy 112 LysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLys 131
    |||
Db 2228 GAGCTGAGGAGATACAAACAGATCTCTTGATGATGAAACCCATTCACAGCGTGGCCACT 2287
    |||
Qy 132 ValCysAsnProAspAsnProGlnGluLysLeuLeuLeuGlnProGlyLysAsnGluIle 151
    |||
Db 2288 GTGTGCCACCCGAAATGGC-----AGTGGCTCAGCTGACGACCAATCTGACGAATGG 2341
    |||
Qy 152 MetAlaAsnSerLeuAspTyrAsnGluArgLeuTrpAlaTrpLysSerTrpArgSerGlu 171
    |||
Db 2342 ATGGCCACATCCCGGAATATGAAACCTGTATGGGATGAGGAGCGGCGGACAGCAG 2401
    |||
Qy 172 ValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGlnMetAla 191
    |||
Db 2402 GCGGGAGAGCCATCTCCATGTTTACCCGAAATAGCTGGAATCTATCAACAGGCTGGC 2461
    |||
Qy 192 ArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGly 211
    |||
Db 2462 CGGCTCAATGGCTATGATGATGACAGGAGACTGTGGAGGCTATGTACAGACACCATCC 2521
    |||
Qy 212 ValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGluAsnThrPheGln 231
    |||
Db 2522 CTGGAG-----CAAGACTGGAGGGGCTCTTCAG 2551
    |||
Qy 232 GluIleLysProLeuTyrGluHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAla 251
    |||
Db 2552 GAGCTGAGGAGCTACTCTCAACCTGCATGCTGAGTGGCGGCGGCTGCACCTGCAC 2611
    |||
Qy 252 Tyr---ProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuGlyAspMet 270
    |||
Db 2612 TACGGGGCCAGACACATCAACCTGAGAGGGGCCCATTCCTGCTCCTGCTGGGGAAATAG 2671
    |||
Qy 271 TrpValArgPheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 290
    |||
Db 2672 TGGGCGGAGACCTGTGTCAACATCTATGACTTGTGTGCTGCTCCCTTCACGCCCTGG 2731
    |||
Qy 291 IleAspValThrAspAlaMetValAspGlnAlaTrpAspAlaGlnArgIlePheLysGlu 310
    |||
Db 2732 ATGACACACACAGAGGCTATGATAAGACAGGCTGAGCGGCCGAGGAGATGTTAAGAG 2791
    |||
Qy 311 AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnIlyPheTrpGluAsn 330
    |||
Db 2792 GGTATGATTTCTTCACTCCCTGGGGCTGCTCCGCTGCTCTGATGATTTCTGGAACAAG 2851
    |||

```

[illegible]

[illegible]







1

Pred. No.: 2.72e-35 Length: 848  
 Score: 381.00 Matches: 87  
 Percent Similarity: 58.54% Conservative: 33  
 Best Local Similarity: 42.44% Mismatches: 63  
 Query Match: 8.88% Indels: 22  
 DB: 4 Gaps: 5

US-09-978-385-2 (1-805) x US-09-247-155-27 (1-848)

```

QY 612 ProTyrAlaAspGlnSerIleValAlaArgIleSerIleValLeuGlyAspIle 631
DB 86 CGAGGTGCAAGAAATGCTTTTAACTGAGACTTACTATCAGACAGCTCGGAGATATA 145
QY 632 AlaTyrGluTyrPasnAspAsnGluMetTyrIleuPheArgSerSerValAlaTyrAlaMet 651
DB 146 GCATTCGCTGGGATACCAATCAATAAATCTTCAAGAGAGGTGCTTCCCATG 205
QY 652 ArgGluTyrPheLeuValIleValAsnGluMetIleuPheGlyGluIleAspValArg 671
DB 206 AGA-----AAAGTTCACACAGAGACACACA---GAAATTTCCCATGTCCTA 250
QY 672 ValAlaAsnLeuValProArgIleSerPheAsnPhePheValThrAlaProIleValAsnVal 691
DB 251 CTTTGCATATGTAACCCAGAGAGGTATCATCTGTGTGTGTACAGACCTTCAAAA--- 307
QY 692 SerAspIleIleProArgThrGluValGluValAlaIleArgMetSerArgSerArgIle 711
DB 308 AATCACCACCTTCCTGCTGCTGTGAGGTGCATACAGCATAGAAAGAACAGACCGGATC 367
QY 712 AsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLeuGlyIleGlnProThrLeu 731
DB 368 AACCATGCTCTTCTTAAATGACCAAACTCTGGAATTTTAAAAATCCCTCCACACTT 427
QY 732 GlyProProAsnGlnProProValSerIleThrIleuIleValPheGlyValIleMetGly 751
DB 428 GCACACCCATGAGACCCCTGCTGTCGCCATCGATATATATATATATGCTATTTTTCG 487
QY 752 ValIleValIleValIleValIleValIlePheThrGlyIleArgAspArgIleValIle 771
DB 488 ATCATCATATGTTGCAATGCTGACTACTGATTTTATCAGAGGATCTGCAACGTAADARAAG 547
QY 772 AsnIleAlaArgSerGly----- 777
DB 548 AACAAAGAACCATCTGAAGTGTGATGACCGCTGAARATAAKTGTGAACATCATGATCAAT 607
QY 778 GluAsn-----ProTyrAlaSerIleAspIleSerIleValGluAsnAsnProGlyPhe 795
DB 608 GAAATGCGATCCCTCTGATCCCTGACATGGAAGGAGGACATATTATATGATGCCCTTC 667
QY 796 GluAsnThrAspAsp 800
DB 668 ATGACAGAGATGAG 682

```

# RESULT 12

US-08-157-171-6

Sequence 6, Application US/08157171

Patent No. 5736323

GENERAL INFORMATION:

APPLICANT: Soubrier, Florent

APPLICANT: Hubert, Christine

APPLICANT: Corvol, Pierre

TITLE OF INVENTION: Agents and Procedures for the Study of

TITLE OF INVENTION: the Genetic Polymorphism of the Angiotensin I Converting

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merchant & Gould

STREET: 90 South 7th Street, 3100 No. 5736323west Center

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/157,171  
 FILING DATE: 24-JAN-1994  
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Hillson, Randall A.

REGISTRATION NUMBER: 31,838

REFERENCE/DOCKET NUMBER: 8076.1030USNO

TELEPHONE: 612-332-5300

TELEFAX: 612-332-9081

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 192 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

IMMEDIATE SOURCE:

CLONE: Exon 18

US-08-157-171-6

## Alignment Scores:

Pred. No.: 4e-07 Length: 192  
 Score: 133.50 Matches: 30  
 Percent Similarity: 56.16% Conservative: 11  
 Best Local Similarity: 41.10% Mismatches: 21  
 Query Match: 3.11% Indels: 11  
 DB: 1 Gaps: 2

US-09-978-385-2 (1-805) x US-08-157-171-6 (1-192)

```

QY 196 TyrGluAspTyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyr 215
DB 3 TATGATGATGACAGGAGGACGCTGAGGTATGTATACGACACCATCCCTGAG----- 56
QY 216 AspTyrSerArgIleLeuIleGluAspValGluIleThrPheGluGluIleLeu 235
DB 57 -----CAAGACCTGAGCGGCTTCCAGAGCTGACAGCA 92
QY 236 LeuTyrGluIleHisAlaTyrValArgAlaIleValMetAsnAlaTyr---Proser 254
DB 93 CTCTACCTCAACCTGATGCTGACGCGCGGCGCTGACACGCTACGACGAGGCGCCAG 152
QY 255 TyrIleSerProIleGlyIleuProAlaHisLeu 267
DB 153 CACATCAACCTGAGGAGGACCATCTCTGCTGACATGCTG 191

```

# RESULT 13

US-08-392-625-16

Sequence 16, Application US/08392625

Patent No. 5837485

GENERAL INFORMATION:

APPLICANT: Entian, Karl-Dieter

APPLICANT: Gtz, Friedrich

APPLICANT: Schnell, No. 5837485bert

APPLICANT: Augustin, Johannes

APPLICANT: Engelke, Gerhard

APPLICANT: Rosenstein, Ralf

APPLICANT: Kaletta, Cortina

APPLICANT: Kleinf, Cora

APPLICANT: Wieland, Bernd

APPLICANT: Kupke, Thomas

APPLICANT: Jung, G nther

APPLICANT: Kellner, Roland

TITLE OF INVENTION: Biosynthetic Process For The Preparation

TITLE OF INVENTION: Of Chemical Compounds

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
 STREET: 1100 New York Avenue  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20005  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 FILING DATE: US/08/392,625  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/876,791  
 FILING DATE: 30-APR-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Esmond, Robert W.  
 REGISTRATION NUMBER: 32,893  
 REFERENCE/DOCKET NUMBER: 0652-0980002  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 371-2600  
 TELEFAX: (202) 371-2540  
 INFORMATION FOR SEQ ID NO: 16:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 8700 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-392-625-16

Alignment Scores:  
 Pred. No.: 0.00153 Length: 8700  
 Score: 127.00 Matches: 152  
 Percent Similarity: 32.58% Conservative: 121  
 Best Local Similarity: 18.14% Mismatches: 261  
 Query Match: 2.96% Indels: 304  
 DB: Gaps: 39

US-09-978-385-2 (1-805) x US-08-392-625-16 (1-8700)

QY 1 MetSerSerSerSerTripleuLeuSerLeuValAlaValThraAlaGlnSerThr 20  
 DB 2562 ATTAGTGAAGCAGCATATATCTCTGCTTATATCTCTATCATTTGGTACAAAACCT 2621  
 QY 21 lIeGlunGlunAlaLysThrPheLeuAspLysPheAsnHISglu- 35  
 DB 2622 ATTAGCAATTATCAGCAATTTTATGATTAATATGATTTGAACACACTAGTAATTTA 2681  
 QY 36 -----AlaGluAspLeu-----PheTyrGlnSerSerLeuAlaSerTripsnTyr 50  
 DB 2682 AAGCAATTCCTCTCAGATTAATTAATGATTTGGCTATCCCAAAAAAGACAGATTATAGTTT 2741  
 QY 51 AsnThrAsnIleThrGluGluAsnValGlnAsnMetAsnAlaGlyAspLysTripsr 70  
 DB 2742 TCTAATAACATT----- 2753  
 QY 71 AlaPheLeuLysGluGlnSerThrLeuAlaGlnMetTyrProLeuGlnGluIleGlnAsn 90  
 DB 2754 GCATTTTAAAGAAAAG----- 2771  
 QY 91 LeuThrValLysLeuGlnIleGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGlu 110  
 DB 2772 -----TATTGCTTGCATTCATAAATAACAGCCATTTGAATTAACAGAA 2816  
 QY 111 AspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGly 130  
 DB 2817 AACGACGTTAAATTTTAAAGAA--AATATACAGTTTCTAAATATCAATCG-- 2867  
 QY 131 LysValCysAsnProAspAsnProGlnGluLysLeuLeuLeuGluProGlyLeuAsnGlu 150

DB 2868 -----CCTGTTCAACTGAA 2882  
 QY 151 lIeMetAlaAsnSerLeuAspTyrAsnGluAlaGluLeuThrPalTrrpGluSerTrrpArgSer 170  
 DB 2883 ATA-----TATAGTGAGATATATTT- 2903  
 QY 171 GluValGlyLysGlnLeuArgProLeuTyrGluGluTyrValValLeuLysAsnGlnMet 190  
 DB 2904 -----GGAATATTCATTAATAAGGT--TATGAGATTTTCCGCTGATTAAGTCAATATTA 2954  
 QY 191 AlaArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrrpArgGlyAspTyrGluValAsn 210  
 DB 2955 GATCTTTTAATGCCGTCGAACCTTTGGAGAGTT--ACGGAAATTTCAATTAATAAG 3011  
 QY 211 GlyValAspGlyTyrAspTyrSerArgGlyGlnLeuIleGluAspValGlnHisThrPhe 230  
 DB 3012 -----AAAAAATTCATTAACAAGAAATTAAGTGTAT- 3044  
 QY 231 GluGluIleLysProLeuTyrGluHisLeuHisAlaTyrValArgAlaLysLeuMetAsn 250  
 DB 3045 -----CATTAACATTAATTCATGAAATGAATGTTTGA 3080  
 QY 251 AlaTyrProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuLeuLysPmet 270  
 DB 3081 ATAGCCCAATTAATGAAGGTCCTTAACCTCAAGAAATGTAAATATTTTGAATAT- 3137  
 QY 271 TrrpGlyArgPheThrPheThrAsnLeuTyrSerLeuThrValProPheGlyGlnAspProAsn 290  
 DB 3138 ---AATAGATATATATATCTTTTAATTTAAATTTTACCT-----AAAAGTGAT 3185  
 QY 291 lIeAspValThrAspAlaMetValAspGlnAlaTrrpAspAlaGlnArgIlePheLysGlu 310  
 DB 3186 ATAGATTAATATGACATATTTATTTGAGCTACATTTAAACAACCTTATCTATTTCTGAA 3245  
 QY 311 -----AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnGly- 326  
 DB 3246 AAGCATGATTCAGAAATTTGATTTGATTAATTCATTAATTAATTAATTAATTAATTAAT 3305  
 QY 327 -----PheTrrpGluAsnSerMetLeuThrAspProGly 337  
 DB 3306 GAATTAACAAATTTTAAAGAAATTTTCAATTGAAAAAACAATTTTATACACTTAA 3365  
 QY 338 Asn-----ValGlnLysAlaVal 343  
 DB 3366 ACTGAAAGAGCATTCATTCATTCCTTTTGTCAAGAAATTAATTTAAATAATATTAAT 3425  
 QY 344 CysHisProThrAlaTrrpAspLeuGlyLysGlyAspPheArgIleLeuMetCysThrLys 363  
 DB 3426 TTAAGACCACTTCTGGAATAATTAATTCGAATGTTTTCGAA-----ACTGAA 3476  
 QY 364 ValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGlnTyrAspMet 383  
 DB 3477 AATGGTTAATATAGGTTGCGCACTATP--AGAGAAAAATGGCATATTCCAAAAGATGTA 3533  
 QY 384 AlaTyrAla----- 386  
 DB 3534 ATTATTCCTTTTGGAGATAATTCATTCATTAATTTAATTAATTAATTAATTAATTAAT 3593  
 QY 387 -----AlaGlnProPheLeuAsnArgAsnGlyAlaAsn 397  
 DB 3594 ATACTAAAAAAGACCTAAAAAAGATGATGAGATTCGATATTTAGCAAGCTTTATCAAT 3653  
 QY 398 GluGlyPheHisGluAlaValGlyGluIleMetSer----- 409  
 DB 3654 GAATCTAATATATGAGAGAAATGTTAGAAATTTGATAGCCATTAATTAATAAAACTAGTTTA 3713  
 QY 410 -----LeuSerAlaAlaThrProLys-----HisLeuLys--- 419  
 DB 3714 AAGAGCAATCTTTTCATTAATCACTAAATAATAGAAATAGCACTTCATTAATTAATTAAT 3773  
 QY 420 -----SerIleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsn 437  
 DB 3774 TGGTTTCAATTCATTTAAGTATTTCTTAATAACATACCAAGAT-----AAT 3818





```

Db 4320 TATGAATTTTAAAAAGAAATTCGAACTTCATGAAATTTCTATTAAATAATAGT 4379
      :|||||:
      :-----ArgThrGluValGluLysAlaIleArgMetSerArgSerArgIle 711
Qy 697 -----:|||||:
      :-----ATTTAGAAATCTTAAAGACACATAAAAAAGCTTAATCTACTTCAGTATGAGATA 4439
Db 4380 ATTTAGAAATCTTAAAGACACATAAAAAAGCTTAATCTACTTCAGTATGAGATA 4439
      :|||||:
Qy 712 AsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLeuGlyIleGlnPro 729
      :|||||:
Db 4440 ATTGCACTTTT-----ATTAACATCGCTTGAATATGAAATATTCGATTAATTCCT 4490
      :|||||:

RESULT 15
US-08-645-193B-18
: Sequence 18, Application US/08645193B
: Patent No. 5962253
: GENERAL INFORMATION:
: APPLICANT: Kupke, Thomas
: APPLICANT: Gotz, Friedrich
: APPLICANT: Kempter, Christoph
: APPLICANT: Jung, Gunther
: TITLE OF INVENTION: Oxidative Decarboxylation of Peptides
: NUMBER OF SEQUENCES: 70
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
: STREET: 1100 New York Avenue, Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/645,193B
: FILING DATE: 13-MAY-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Esmund, Robert W.
: REGISTRATION NUMBER: 32,893
: REFERENCE/DOCKET NUMBER: 0652,1540000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 8700 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: both
: MOLECULE TYPE: cDNA
: US-08-645-193B-18

Alignment Scores:
Pred. No.: 0.00454 Length: 8700
Score: 123.00 Matches: 153
Percent Similarity: 32.46% Conservative: 120
Best Local Similarity: 18.19% Mismatches: 258
Query Match: 2.87% Indels: 310
Gaps: 40

US-09-978-385-2 (1-805) x US-08-645-193B-18 (1-8700)
Qy 1 MetSerSerSerTyrPheLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20
      :|||||:
Db 2662 ATTTAGTGAAGCATATATCTCTGTTATTAATCTCTAATCATTTTGGTACAAAACAT 2621
      :|||||:
Qy 21 IleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGlu----- 35
      :|||||:
Db 2622 ATTTAGCAATTTTTCACGAATTTTATGATTAATATGATTTGAACAACTAGTAATTTTA 2681
      :|||||:

```

```

Qy 36 -----AlaGluAspLeu-----PheTyrGlnSerSerLeuAlaSerTyrPasn 50
      :|||||:
Db 2682 AACCAATTCCTCTCACATATTAATGATTTGGCTATCCCAAAAAGACGTTATAGTTT 2741
      :|||||:
Qy 51 AsnThrAsnIleThrGluGluAsnValGlnAsnMetAsnAlaGlyAspLysTyrSer 70
      :|||||:
Db 2742 TCTAATACAT----- 2753
      :|||||:
Qy 71 AlaPheLeuLysGluGlnSerThrLeuAlaGlnMetTyrProLeuGlnIleGlnAsn 90
      :|||||:
Db 2754 GCATTTTAAAGAAAG----- 2771
      :|||||:
Qy 91 LeuThrValLysLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGlu 110
      :|||||:
Db 2772 -----TATTGCTTGCATATTAATAATACAGCCATTTGAATTAACAGCA 2816
      :|||||:
Qy 111 AspLysSerLysArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGly 130
      :|||||:
Db 2817 AACGACGTTAAATAATTTGAAAG-----AATTAATACGTTTCAATAATCAATCG----- 2867
      :|||||:
Qy 131 LysValCysAsnProAspAsnProGlnGluCysLeuLeuLeuGluProGlyLeuAsnGlu 150
      :|||||:
Db 2868 ----- 2882
      :|||||:
Qy 151 IleMetAlaAsnSerLeuAspTyrAsnGluArgLeuTyrPheValLeuLysAsnGluMet 170
      :|||||:
Db 2883 ATA-----TATAGTCGATATATTT----- 2903
      :|||||:
Qy 171 GluValGlyLysGlnLeuArgProLeuTyrGluGlyTyrValValLeuLysAsnGluMet 190
      :|||||:
Db 2904 -----GGAATTCATTAATAAGGT-----TATGAGATTTGGCGTGATAGTCAATATTA 2954
      :|||||:
Qy 191 AlaArgAlaAsnHisTyrLysAspTyrGlyAspTyrTyrPheGlyAspTyrGluValAsn 210
      :|||||:
Db 2955 GGATCTTTTAAGCCCGCGTCAACTTTGGAAGCTTT--ACGGAAATTTCAATTAAG 3011
      :|||||:
Qy 211 GlyValAspGlyTyrAspTyrSerArgGlyGlnIleGlnAsnValGlnHisThrPhe 230
      :|||||:
Db 3012 -----AAAAAATCAATTAACAAAAGAAATAGTGCAT----- 3044
      :|||||:
Qy 231 GluGluIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsn 250
      :|||||:
Db 3045 -----CATTCATTAATTAATCAATGAAGAATGTTAGAA 3080
      :|||||:
Qy 251 AlaTyrProSerTyrTyrLeuSerProIleGlyCysLeuProAlaHisLeuGluLysPmet 270
      :|||||:
Db 3081 ATTAAGCAATTAATAAGTCACTCTTAATCAAGAAATGTAATAATTTGAATAT-- 3137
      :|||||:
Qy 271 TyrGlyArgPheThrPheThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsn 290
      :|||||:
Db 3138 -----AATGAATATATATCTGTTAAATTTAATTTACG-----AAAAGTAT 3185
      :|||||:
Qy 291 IleAspValThrAspAlaMetValAspGlnAlaTyrPaspAlaGlnArgIlePheLysGlu 310
      :|||||:
Db 3186 ATGAGATTAATGAACATATTAATGAGCTACATTAACAACTTATCTATATCTGAA 3245
      :|||||:
Qy 311 -----AlaGluLysPhePheValSerValGlyLeuProAsnMetThrGlnIle 326
      :|||||:
Db 3246 AACATGATTCAGAAATGATTTGATCTTAATTAATGATTAATGATGAGTTGGAGCT 3305
      :|||||:
Qy 327 -----PheTyrGluAsnSerMetLeuThrAspProGly 337
      :|||||:
Db 3306 GAATTTACAAATTTTAAAGAGAAATTTGATTTGAAAAACAAATTTATACAACTATA 3365
      :|||||:
Qy 338 Asn-----ValGlnLysAlaVal 343
      :|||||:
Db 3366 ACTGAAGAGGATGACTCATTAATCTTTTGTCCAAAGATTTATTAATAAATTTATTT 3425
      :|||||:
Qy 344 CysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIleLeuMetCysThrLys 363
      :|||||:
Db 3426 TTAATAACGAGCTTCTGGAATAATTAATTCAGAAATGTTTTCGAA-----ACTGAA 3476
      :|||||:
Qy 364 ValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGlnTyrAspMet 383
      :|||||:

```

```

Db 3477 AATGTTAAATAGTTCGCAACTATT--AGAGAAAAATGGCATATTCCAAAAAGATGTA 3533
QY 384 AlAtyRaA----- 386
Db 3534 ATTATGCTTTGGAGATATCCATTCCTAATTAATTATTAATGACAGCATCTCAT 3593
QY 387 -----AlaGlnProPheLeuLeuArgAsnGlyAlaAsn 397
Db 3594 ATACTAAAAAGAACTAAAAAACATGATGAGTTCGATATTCGAAACCTTATCAAT 3653
QY 398 GlnGlyPheHisGluAlaValGlyGluIleMetSer----- 409
Db 3654 GAATCTAATATGAGAGAAATGTTAGAAATGTTACGCCATTTATATATATATATATAT 3713
QY 410 -----LeuSerAlaAlaThrProLys-----HisLeuLys--- 419
Db 3714 AAAGAACAACTTTTCATTAATACCTAAAAAATAGAAATAGACATTCATATATCTTAAAGAT 3773
QY 420 -----SerIleGlyLeuLeuSerProAspPheGlnGlnAspAsnGlnThrGluLeuAsn 437
Db 3774 TGGTTTCAATTCATTTAATATTCCTTAAACATACCAACAT-----AAT 3818
QY 438 PheLeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheThrIleMetLeuGlu 457
Db 3819 TTTATTCAGATTATCTATTAATACCATTTATACGGAATTTAAAGTTAATATATTTATTAAT 3878
QY 458 LysThrArgThrMetValPheLysGlyGluIleProLysAspGlnThrMetLysLysThr 477
Db 3879 AAATTTTTTACATAAAATTTAAAGAGAT-----GAAATTTTATAAAA----- 3923
QY 478 TrpGluMetLysArgGluIleValGlyValGluProValProHisAspGlnThrIle 497
Db 3924 TTAAGATTATTAAGAGA-----GATGAAGATTAT 3953
QY 498 CysAspProAlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyr----- 516
Db 3954 -----TCCAATTATTAATCTTCAATTAATAAAATTTGAAA 3986
QY 517 -----ThrArgThrLeuTyrGlnPheGlnPheGlnIleLeuLysGln 531
Db 3987 GATTATTCCTTAAATAGTGAATTAATGACTATTAATGACTATTAATGCTATGCTCTGAA 4046
QY 532 AlAlaLysHisGlnGlyProLeuHisLysCysAspIleSerAsnSerThrGluAlaGly 551
Db 4047 GATATATAGATATGCTGACACGATATTAAGATATGAGAAATTTTATGATGAT 4106
QY 552 GlnLysLeuPheAsnMetLeuArgLeuGlyLysSerGlnProThrPheLeuAlaLeuGlu 571
Db 4107 AGTCTATTA-----TCAATTAATATTAATCAATCAAG--TTCAAAATTTCCAAAAAGAA 4157
QY 572 AsnValValGlyAlaLysAsnMetAsnValArgProLeuLeuAsnThrPheGluProLeu 591
Db 4158 TTTATGCTTCT-----ATATCAATATGATTTTATTTATGATTAATTTAGAA----- 4202
QY 592 PheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGlyTrpSerThrAspTrpSer 611
Db 4202 ----- 4202
QY 612 ProTyrAlaAspGlnSerIleLysValArgIleSerLeuLysSerAlaLeuGlyAspLys 631
Db 4203 -----ATTAAATTAAGATAGAGAAAGAAATTTTAAATTAATTAATGCG-----GAAAGAT 4250
QY 632 AlaTyrGlnTrpAsnAspAsnGlnMetTyrLeuPheArgSerSerValAlaTyrAlaMet 651
Db 4251 TTTATTCGATGATATGAC-----ATA 4271
QY 652 ArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPheGlyGlnGluAspValArg 671
Db 4272 AGAGAAATAT----- 4280
QY 672 ValAlaAsnLeuLysProArgIleSerPheAsnPhePheValThrAlaProLysAsnVal 691

```

```

Db 4281 -----AAAAATTATAGCTAAACTT-----ACCAATCTAAAAATGAC 4319
QY 692 SerAspIleIlePro----- 696
Db 4320 TATGAATTTTAAAAAGAAATTTCCGAATCTTCATGATTTCTATTATTAATAAATTACT 4379
QY 697 -----ArgThrGluValGlnLysAlaIleArgMetSerArgSerArgIle 711
Db 4380 ATTTTGAATAATCTTAAAAAGACACTACAAAAAACCCTATATATCTTACAGTTCTAGATA 4439
QY 712 AsnAspAlaPhe-----ArgLeuAsnAspAsnSerLeuGlnPheLeuGlyIleGln 728
Db 4440 ATTTGACGATTTATACATGCGTTGTAAT-----AGAAATTCGATTAAT 4487
QY 729 Pro 729
Db 4488 CCT 4490

RESULT 16
US-09-280-116-114/C
; Sequence 114, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 114
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: angiotensin-converting enzyme
; NAME/KEY: misc.feature
; LOCATION: (1)-(789)
; OTHER INFORMATION: n = a, t, c or g
US-09-280-116-114

Alignment Scores:
Pred. No.: 0.000133 Length: 789
Score: 121.00 Matches: 31
Percent Similarity: 47.78% Conservative: 12
Best Local Similarity: 34.44% Mismatches: 29
Query Match: 2.82% Indels: 18
Gaps: 3

US-09-978-385-2 (1-805) x US-09-280-116-114 (1-789)
QY 186 LeuLysAsnGlnMetAlaArgAla-----AsnHisTyrGluAsp 198
Db 384 CTAAGGCGTGAAGATGCTGACCTCTGAGTCTCTCTCTGCTAGGTACAAAGAC 325
QY 199 TyrGlyAspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSer 218
Db 324 ATGGGGCCCTTGCGCACTCCAAATATGATGCGGATGCCGAG----- 280
QY 219 ArgGlyGlnLeuIleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlu 238
Db 279 -----CAAGACCTTGAGCGCGGTATTTCCAGAGAGCTGCGGCACTCTACCTG 235
QY 239 HisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyr--ProSerTyrIleSer 257
Db 234 AACCCGACACCTAGCTGCGGAGGCGCTCCACCGCCACATATGAGGCCGAGCATCATGAC 175
QY 258 ProIleGlyCysLeuProAlaHisLeuLeu 267
Db 174 CTGAGGGGGCCCATCCCTGCGCCACTCTCTG 145

RESULT 17

```



```

US-09-453-702B-1/c
; Sequence 1, Application US/09453702B
; Patent No. 6365723
;
GENERAL INFORMATION:
;
APPLICANT: Blatner, Frederick R.
;
Burland, Valerie
;
Perna, Nicole T.
;
Plunkett, Guy
;
Welch, Rod
;
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
;
NUMBER OF SEQUENCES: 265
;
CORRESPONDENCE ADDRESSES:
;
ADDRESSEE: Quarles & Brady
;
STREET: 1 South Plunkney Street
;
CITY: Madison
;
STATE: WI
;
COUNTRY: US
;
ZIP: 53701-2113
;
COMPUTER READABLE FORM:
;
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
;
COMPUTER: IBM PC compatible
;
OPERATING SYSTEM: PC-DOS/MS-DOS
;
SOFTWARE: Word Perfect 8.0
;
CURRENT APPLICATION DATA:
;
APPLICATION NUMBER: US/09/453,702B
;
FILING DATE: 03-Dec-1999
;
CLASSIFICATION: <Unknown>
;
PRIOR APPLICATION DATA:
;
APPLICATION NUMBER: 60/110,955
;
FILING DATE: 04-DEC-1998
;
ATTORNEY/AGENT INFORMATION:
;
NAME: Seay, Nicholas J.
;
REGISTRATION NUMBER: 27386
;
REFERENCE/DOCKET NUMBER: 960296.95017
;
TELECOMMUNICATION INFORMATION:
;
TELEPHONE: (608) 251-5000
;
TELEFAX: (608) 251-9166
;
INFORMATION FOR SEQ ID NO: 1:
;
SEQUENCE CHARACTERISTICS:
;
LENGTH: 6506
;
TYPE: nucleic acid
;
STRANDEDNESS: double
;
TOPOLOGY: linear
;
MOLECULE TYPE: DNA (genomic)
;
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
;
US-09-453-702B-1
;
Alignment Scores:
;
Pred. No.: 0.00548
;
Score: 120.50 Length: 6506
;
Percent Similarity: 32.57% Matches: 186
;
Best Local Similarity: 19.29% Mismatches: 128
;
Query Match: 2.81% Indels: 323
;
GB: 4 Gaps: 51
;
US-09-978-385-2 (1-805) x US-09-453-702B-1 (1-6506)
;
OY 5 SerTrrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThrIleGlnGln 24
;
Db 3778 ACATTATATCTCACACCCGCAAGTAATCGTGTCGACGAGTGCTTGGCCACAACCTTAAA 3719
;
OY 25 AlAlayStrhrheLeuaspIlySpheasnhISgIuAlaGluaspLeuPhetYrGlnSer 44
;
Db 3718 ATTAGAAAAATGCCAACAACTTTGGCCCTCACAAATAAGAAAGTTATTTTAT----- 3668
;
OY 45 LeuAlaSerTrpAnsTrpAnsThrAnsIleThrGlnGluAsnValGlnAsnMetAsn 64
;
Db 3667 -----CTAAATGTTCGGATATCCCTCCAAACAAATCCACAAAT 3629
;
OY 65 -AlaGly--AspIlyStrpSerAlaPheLeuIlySgIuGlnSerThrLeu----- 79
;
Db 3628 GCAGGGCAAAACAAATTTAACTTGATTCGCAAAACCGATTAAGTTACTTTGGCGTCGA 3569

```

Qy	80	AlaIleuIeTyrProLeu-----GlnGluIleGlnAsnIleThrValIleGlnIleu	97
Db	3568	AGCGAATTAACCTCCGGTACATAAAAAAACCCCTTTCTCAATTAAATATCAAAAAACAAGAAAT	3500
Qy	98	GlnAlaLeuGlnIleAsnGlySerValIleuSerGlnAspIleSerylValGlnLeuAsn	117
Db	3508	AACGCCATCTCATTAACAAACAAGAACAGCC-----AAC	347
Qy	118	ThrIleuAsnSrlMetSerThrIleTyrSertThrGlyValCysAsnProAspAsn	137
Db	3475	TGGATA-----ACGGTTACAACTATC-----AAGCACAAAACGTGAAGTT	343
Qy	138	ProGlnIleuGlyLeuIleuIleuGlnPro--GlyLeuAsnGlnIle--MetAlaAsnSer	155
Db	3433	AATAATGAAGAAGTACTTACTGCGCCCTTTGCGCATACGACATTAATTAATAAAATAC	337
Qy	156	LeuAspTyrAsnIleuArgIleuThrAlaIleProIleuSertThrValSerGlnIleValIleGlyGln	175
Db	3373	CACGCATCAAGATATGCAATTAACGGTAGTATATGACGGGAATTAACATCACTTACATAA	331
Qy	176	LeuArgProLeuTyrGlnGluIleuTyrValValIleuIleuSAsnGlu-----MetAlaArgAla	193
Db	3313	ATCGCGCGAGATTAACAGCAAAAAAGTATGTTAAGAGACACACATGGCGCTGCTATT	325
Qy	194	AsnHisTyrGlnAspTyrGlnIleAspTyrThrArgIleValSerylGlu-----Val	209
Db	3253	ATTTTGTGTATTCCTTTCCGGGTATGCTGAAGAACTTTTGATACGACTTATGATC	319
Qy	210	AsnGlnValAspGluTyrAspTyrSertArgIleGlnIleuIleGlnAspValGlnHisThr	229
Db	3193	GGCGCGATGAAGAGAGAAAAAGTTCC-----GAATTCAT	315
Qy	230	PheGluGluIleIleValProLeu-----TyrGlnHisLeuHisAlaTyrVal	244
Db	3157	TTTCGATTAATTAACCAACCGCTTCCAGAAATACGAA--CTGATTTTATGTAATTAAC	310
Qy	245	-----ArgAlaIleValLeuMetAsnAlaTyrProSerylThrIleSerProIle--GlyCys	261
Db	3100	CAGTCGGCGTGAAGAAACAAGATATACAGATCCCAAG-----TCACCAGTCAAAACCACT	3047
Qy	262	LeuProAlaHisIleu-----LeuGlnAspMetThrGluArgPheThrIleAsn	277
Db	3046	CTGCCTAAGGTGCTACTAACCAACTGGGGGTAAAGACCGGTAACTCAATACAGAAGAT	2987
Qy	278	-----IleTyrSerLeuThrValProPheGlyGln-----	287
Db	2986	AATTGATATTTCAGTACGATTAAGCGGTTTCATGTCGACCAATATATCATCGGATATCAGTAA	2927
Qy	288	--IysProAsnIleAspValThrAspAlaMetValAsp-----	299
Db	2926	CACGCACATAAATTAACAGTACACAGCGCATATATCATCAATGAACTGAAGAGAGTACGT	2867
Qy	300	-----GlnAlaIleProAspArgIleGlnArgIlePheIleGlnAlaGlnIleSerylPheValSer	317
Db	2866	CCCCCGGAAGTTGGGAT-----CGCGCATTTGATGACTTTTATACCTCC	2822
Qy	318	ValIleGlyLeu-----	320
Db	2821	TATATTTTAAGTCAGTATCCCTCTTATGACACGACAAATAATAACAGTAATACAGCCAGTTAT	2762
Qy	320	-----	320
Db	2761	GGCGGATTTTAACAGTGAATTAATTTATTTAGTTGCGACTTACACTCGATGCCAGCTAT	2702
Qy	321	-----ProAsnMetThrGlnGlyPheThrGlu--AsnSerMetLeuThrAspProGly	337
Db	2701	AGCAAAACCCGATGATATGAAGAGCACATGGAAGTAATATACCTTTTACCTGGACATGGC	2642
Qy	338	-----AsnValGlnIleValIleValCysHis	345
Db	2641	TGTCGCGAGATATTAACACAGCTTGAGTANGAGAAATATACATCTCCTCATATATTC	2582
Qy	346	ProThrIleAlaTrpAspLeuGlyIleGlyAspPheArgIleIleuMetCysThrIleValThr	365

```

Db 2581 GACCTCTGAGATTAGCGGAATACGTTATTCGGATATGCAATGCTTACCTGATTCCT 2522
OY 366 Metaspasph-----Leuthralahis 373
Db 2521 ATGCGAGGCTTTACACCTTATAGTACAGGCTTGCGCAAGTAAGCTTGAACGGTT 2462
OY 374 Hislumelglnhisileglntryaspmetalaityralalaglnpropheuleuarg 393
Db 2461 TCGCAAAATGGCTACATCATTTACCAAAAAGAGTTCGCCCGGACCATTTACTATCGCA 2402
OY 394 Asncllyalaasnclgly-----Phehisgluvalalagly 405
Db 2401 GATTTCGACACTCTCGCGACCGGCTCTGATCTCGATGTCAGCTAAAGACAGATGCGC 2342
OY 406 Gluilemetserleu-----SerAlaIatThrProlyshisleuysSerilegly 422
Db 2341 AGTGTTCCTTCTTTTGGTGCCCTACTCTCCGTGCTAACAATGCTGCAACCT---GGT 2285
OY 423 LeuLeuSerProasphe-----GlnGluasp 431
Db 2284 ATTCAATTCGATTTTATTCGCGGACGACGAGTAAATATGCTGTAAGAATCAGAG 2225
OY 432 Asnclutlrghuileasnphelu-----LeuLysgluAlaLeuThrileValagly--- 448
Db 2224 GACCTTTCGAGCAAAATATATATATATGAGCTCAACAATCTTTTGACACTTATGCGCGT 2165
OY 449 ---Thrleu-PropheThrlyrmetleuclglystrpargrtrmetvalphelysgllyl 467
Db 2164 ACGATTTTATCCGATTAATGCTAATGCTAATGCTG---2128
OY 467 uileprolyaspclntrpmetlyslstrprrpgleumelylsargluilevalglya 487
Db 2127 -----GAAATGGTTGGAAAT 2114
OY 487 lvalgluprovalpro-----HisaspgluThrlycysasproalase 502
Db 2113 ACACCGCTGGGTCCATCTCTTTGAGCGTACAGATCGAGTGAAGTA---AATATAGA 2055
OY 502 rleuPhehisValSerhasnasptryr-----SerPheIlegrtyrtyrThrarghile 520
Db 2054 TATTAACACAGCAAGGTACAGCTATACAGTCCCTATATATATATCTAGTAAACGCG 1995
OY 520 uTyrlinPhegin-----Pheginl 527
Db 1994 CACTCGATTAGCGTGGCCGCTGCTTATGCTTCGACAGATTACAGACATTCAGCA 1935
OY 527 uAlaLeuGysgluAlaAlaAlaLysHisgluglyProleuHisLysCysaspIleSerans 547
Db 1934 CCATCTTTACGAAATGATAAATAAT---CAGCAGAGTACATATGACGATTT 1884
OY 547 rThrgluAlaglylnLysleuPhehasnmetleuargleuglyLyssergluprotprh 567
Db 1883 TTATGATATTTGGCAGAAAA---AATAGCCTTCTCCCAATTCATGCACT---1835
OY 567 rleuAlaLeuGlnuasnValalaglyAlaLysasnmetasnValarproleu---leuas 586
Db 1834 -----TTATCCATATATCTGGGA---AATGATCATTAAGTGGCTTGGCGGAA 1788
OY 586 nTyrrPhegluproleuphethrtrpleuLysaspLlnasnLysasnserphe---Valgl 605
Db 1787 TTAC-----TGCGGCGCAAGTGAATGCTAAAGATTAACCAATTACG 1746
OY 605 yrrpserthrapptrpserprotyrAlaAspLlnSerileLysValArgileserleuly 625
Db 1745 TTATTCCAATACCTGGCA---CCATCACT---1718
OY 625 sSerAlaLeuGlyaspLysAlaTyrgluTrpAsnaspasnclumetyrleuphearSe 645
Db 1717 -----TATACTTCTCTGC 1704
OY 645 rSerValAlaTyrgluMetarglnTyrrPheleuLysValLysasnclumetileleuph 665

```

```

Db 1703 CAGCAATCCTAT-----1691
OY 665 egllygluaspvalargvalAlaAlaasnleuLysproargIleSerPheasnphelpheva 685
Db 1690 -GATGAAMCGAC-----AAGAGAGAGCGCTTTTAACTGTTAT 1650
OY 685 lThrAlaProlyasnValSeraspIlelleProargThrGluValGluLysAlaIleAr 705
Db 1649 CTCATTCCTCTTCTACTGGGGGATATATTCGCCAAACACGTCACCAAAATTAATCTATC 1590
OY 705 gmetSerargSerargIleasnspAlaPheargleuasnaspasnSerleuGluPhele 725
Db 1589 GAATTCGACCTCATTTCCAAAGATGCTATTCCTCCCAACAATGACT-----1544
OY 725 uGlylleclnProthrleuglyProproasnGlnProprovalSerIletrpleuIleVa 745
Db 1543 -GGATTTACGCAATGACCGGTGAGACATGATCG-----TTAAA 1506
OY 745 lPheglValvalmetcllyalllevalValaglyIlevalIleleuIlePheThrlyl 765
Db 1505 TTATGTAATAT---GT 1491
OY 765 eargaspargLysLysLysasnLysAlaArgSerGlyGluasn-----Pr 780
Db 1490 TAATCAGCAACAAACAAATATATGATACATCGCTTGACAGATTTAACTGAGACATCC 1431
OY 780 oTyrrAlaSerIleaspIleSerLysgllyLysasnaspProglyPhegluasnThrAsps 800
Db 1430 CATCGCCACAAATATATGCGCTATGACCATTTCTAAACACGCTGCAAGAGTGGGAAG 1371
OY 800 pvalGlnThr 803
Db 1370 TATATAGAT 1361

RESULT 18
US-08-961-527-45/c
; Sequence 45, Application US/08961527
; Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 11384 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

```

TOPOLOGY: linear  
 US-08-961-527-45  
 Alignment Scores:  
 Pred. No.: 0.0142  
 Score: 120.50  
 Percent Similarity: 31.288  
 Best Local Similarity: 17.988  
 Query Match: 2.81%  
 Gaps: 45  
 Length: 11384  
 Matches: 146  
 Conserved: 108  
 Mismatches: 238  
 Indels: 321  
 US-09-978-385-2 (1-805) x US-08-961-527-45 (1-11384)  
 QY 23 GluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaLysPheLeuGln 42  
 DB 11096 GAAGCATTAATAAACCTTTATG-----ACGGGTAATAATTTTATCTCCAA 11052  
 QY 43 SerSerLeuLaser----- 47  
 DB 11051 CATTATCTAGGAGCACATAGGAGAACTAAATGAGACATGCTATACCTTCCTGTT 10992  
 QY 48 TrpAsnTyrAsn-----ThrAsnIleThrGluGlu 57  
 DB 10991 TGGGACACCTAATGCTCAGGCTGTCTGCTGCTGATTTACCACTGGATTGAAT 10932  
 QY 58 AsnValGlnAsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSer 77  
 DB 10931 CAGATTCCAAATGATAGAAATGATTTGGGCTGCGAAGCTTT-----ACC 10884  
 QY 78 ThrLeuAlaGln-----MetYrProLeuGlnIleGlnAsnLeuThrValLys 94  
 DB 10883 AATATGCTCCAGAGAGGATTTTAC----- 10857  
 QY 95 LeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerSerValLeuSerGluAspLysSerLys 114  
 DB 10856 ---AATATCATGCTCAGCAGTCAAAATGCTATCAACGATGAGAG- 10815  
 QY 115 ArgLeuAsnThrIleLeuAsnThrMetSerThrIleTyrSerThrGlyLysValCysAsn 134  
 DB 10814 -----ATTGACCTTTTGTCTGCTGAGTATGAGCT- 10785  
 QY 135 ProAspAsnProGlnGluLysLeuLeuGlnIleProGlyLeuAsnGlnIleMetAlaAsn 154  
 DB 10784 -----CGTCCAGGAGACAGGGGCAATGTAACA- 10758  
 QY 155 SerLeuAspTyrAsnGlnLysLeuTrp-----AlaTrpGlnSerTrpArgSerGlu 171  
 DB 10757 -----GAGCTTCCTGAGAGAAATGAGAGATGAGACTTTGGCTGGCAGAGAAACGT 10704  
 QY 172 ValGlyLysGlnLeuArgProLeuTyrGluGlyValValLeuLysAsnGluMetAla 191  
 DB 10703 TGGGGCTTTGAGAGAGCGCTGCTC----- 10680  
 QY 192 ArgAlaAsnHisTyrGlnAspTyrGlnLysAspTyrTrpArgGlyAspTyrGlnValAsnGly 211  
 DB 10679 -----AATATTATGAGAGTTCACCGCTGATCATGAGAA- 10641  
 QY 212 ValAspGlyTyrAspTyrSerArgGlyLysLeuLeuIleGlnLysValGlnHisThrPheGlu 231  
 DB 10640 TCTGATGGCAGCTTATATGTTGGCCAGCTCAAGATGATCATCTCTATCTCGTT 10581  
 QY 232 GlnIleLysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAla 251  
 DB 10580 GAAATGAAC-----TATATCTCATTTGAG----- 10557  
 QY 232 TyrProSerTyrIleSerProIleGlyCysLeuProAlaHisLeuGlnLysAspMetTrp 271  
 DB 10556 -----TTTATGCC-----TTGATGTCCATCTTTGGGCTTGAGTTGG 10518  
 QY 272 GlyArgPheTrpTrpAsnLeuTyrSerLeuThrValProPheGlyLysLysPro----- 289  
 DB 10517 GGGATACAGCTTATGGGTTACTCTGCTTTAGAGCATCTTATGGCGACAGAGAGTTT 10458

QY 290 -----AsnIleAspValThrAspAlaMetValAsp 299  
 DB 10457 CAGATTTCGAGAGAGTGTCAATACCAATAATATGCGTT-----ATTGTGAC 10407  
 QY 300 GlnAlaTrpAspAlaGlnArgIlePheLysGluAlaGlnLysPhePheValSerValLys 319  
 DB 10406 -----TGGGTACCACTGACT 10392  
 QY 320 LeuProAsn-----MetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGly 337  
 DB 10391 TTACCATCAACAGATGATGCTTGA-CCCTATTTATGATGGACACCGACACTTTGAATACCA 10333  
 QY 338 AsnValGlnLysAlaValLysHisProThrAlaTrp-----AspLeuGly 352  
 DB 10332 GACCATATATAGGCT-----CATACCATGTTGGGCTGCCCTTAATTTTGACCTTGA 10279  
 QY 353 LysGlyAspPheArg-----IleLeuMetCysThrLysValThrMetAspPhe--- 369  
 DB 10278 AAAAATGAAGTCCAGTCCTTTAATTTCTTCAATTAGCATGATGATGATGAT 10219  
 QY 370 -----LeuThrAlaHisHisGluMetGlyHisIleGlnTyrAspMetAla 384  
 DB 10218 TTGATGATGATTTGCTGATGCTGTTAGCAACATGCTCTATTTGACTATGATGATGCT 10159  
 QY 385 TyrAlaAlaGlnProPheLeu---LeuArgAsnGlyAlaAsn-----GluGlyPhe 400  
 DB 10158 -----CATGACACCTTAATTAAGATGCGGAGAAATCTCACTAATGAAGTTAT 10111  
 QY 401 His-----GlnAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeu 418  
 DB 10110 TATTTCTTCAGCGCTTGAATGAGGTTATTAAGTTAGAA----- 10072  
 QY 419 LysSerIleGlyLeuLeuSerProAspPheGlnGlnAspAsnGlnThrGluIleAsnPhe 438  
 DB 10071 -----TATCAAGATGTCATGATGATGCGAAGAAAGTTCGTGCGC 10030  
 QY 439 LeuLeuLysGlnAlaLeuThrIleValGlyThrLeuProPheTrpTyrMetLeuGlnLys 458  
 DB 10029 ATCAAGATTACGGGATGAGATGAAGAGATGCTGCTCAGGATTTGACTAC-----AAA 9979  
 QY 459 TrpArg-----TyrMet-----ValPheLysGlyLysIlePro----- 469  
 DB 9978 TGGACATGGCTGGATGATGATATTCCTCGCTTCTCAAGAAAGATCCGATCATGCT 9919  
 QY 470 -----LysAspGlnTrpMet 474  
 DB 9918 AATATGACTTTAACCTGCTGACTTTCAGCTTATGATGTTTTCAGAGAAATATATCTC 9859  
 QY 475 LysLysTrpTrpGluMetLysArgGluIleValGlyValValGluProValProHisAsp 494  
 DB 9858 TTG-----CCATTCCTGCGACGAT 9841  
 QY 495 GluThrTyrCysAspProAlaSerLeuPheHisValSer-----AsnAsp 509  
 DB 9840 GAAGTGGTTATGCGCAAGAGATATGATGATGATGATGATGATGATGATGATGATGAT 9781  
 QY 510 TyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPheGlnGlnAlaLeu 529  
 DB 9780 TTGCGAGGCTTG-----CGCAATCTTATACGATCAAA-----ATT 9745  
 QY 530 CysGlnAlaAlaLysHis----- 535  
 DB 9744 TGTCACTCTGTAAGAAATTCCTCTTATGATGATGATGATGATGATGATGATGATGATG 9685  
 QY 536 -----GluGlyProLeu----- 539  
 DB 9684 AATCTGAAGACAGTTGATGATGCTTACCTAGAACACCAATGATGATGATGATGATGATG 9625  
 QY 540 -----HisLysGly----- 542  
 DB 9624 TATTTGCTTCTCAGCTTAACCACTTTTACAAGATCATGCTGCTGCTGCGAAATGAT 9565  
 QY 543 -----AspIleSerAsnSerThrGlnAlaGlnLysLeuPheAsn 556



Query Match: 2.808 Indels: 298  
 DB: 4 Gaps: 42  
 US-09-978-385-2 (1-805) x US-08-961-527-115 (1-11303)

QY 40 PheTyrGlnSerSerLeuAlaSerTrpAsnThr--AsnIleThrGluAsn 58  
 DB 2860 TTTTTCATGACTATATTTGTCATGATATTAACAGATGAGACATTTGTCATGAGATA 2801  
 QY 59 ValGlnAsnMetAsnAlaGlyAspLysTrpSerAlaPheLeu--LysGlu 75  
 DB 2800 ATGAAATGCGATCAAAAATG--GTGCGTCTGTTTATGCTACATAAAAGAG 2747  
 QY 76 GlnSerThrLeu-----AlaGlnMetTyrProLeuGlnIleGlnAsnLeu 91  
 DB 2746 GAAATCAATTTGAAATTTTAAAGGAGAGTTTAT--CCAACTCTGTATTA 2696  
 QY 92 ThrValLysLeuGlnLeuGlnAlaLeuGlnGlnAsnGlySerValLeuSerGluAsp 111  
 DB 2695 ACAGACAAAGCTAGTAGGTAGATGACTCTCAACTGGAAGT--TTTGAGGAT 2645  
 QY 112 LysSerLysArgLeu----- 116  
 DB 2644 AGAAGCGACACACTATCTATATATAGAGATTGCGCAATTTGCTAGATTATATGAG 2585  
 QY 117 --AsnThrIleAsnThrMetSerThrIle-----TyrSerThrGlyLys 131  
 DB 2584 ACATCAAAACCTACTAGATATATATGACGACATTTTTCATCTGCACCTTAATAAAGGAGA 2525  
 QY 132 ValLysAsnProAspAsnProGlnGluLysLeuLeuGluProGlyLeuAsnGluIle 151  
 DB 2524 TTT--AACCGCCGAAAAT--TTATTTATTA-----GAATTA 2495  
 QY 152 MetLAsnSerLeuAspTyrAsnGluArgLeuThrPalatrlatrlgIuserTrpArgSerGlu 171  
 DB 2494 AAAGCAATTTTGGCGTATATAGTAGTACGCTGATTTTGTGTAATCTATGAAACCTTA 2435  
 QY 172 ValGlyLysGlnLeuArgProLeu-----TyrGluGluTyr 183  
 DB 2434 AAAAGA--ACAGCTCGACCTGGATAGAGACAGATGAGCAATCTGCTTAGAGAGATGGA 2338  
 QY 184 ValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGlyAspTyrTrp 203  
 DB 2377 ATTATTTAGCCGAATGATAGCTTATGCGAT--CATCTCAAGGATTTATTTGG 2324  
 QY 204 --ArgGlyAspTyr-----GluValAsnGlyValAsp----- 213  
 DB 2323 GATGACACATCTGTTTATTTGAGAGAGAAAATCAAAATGACCTGATTTCTTTGCCCTAT 2264  
 QY 214 GlnTyrAspTyrSerArgGlyGlnLeuIleGluAspValGlnHisThrPheGluGluIle 233  
 DB 2263 GGGGATGATTTAT--AGAGGA--GCAATCGAGATTTTTCACAT----- 2225  
 QY 234 LysProLeuTyrGlnHisLeuHisAlaTyrValArgAlaLysLeuMetAsnAlaTyrPro 253  
 DB 2224 ----TTGACTGGTTCAACACC 2207  
 QY 254 SerTyrIleSerProIleGlyCysLeuProAlaHisLeuGluAspMetTrpGlyArg 273  
 DB 2206 ----TTGTTGCCAAGATATGCTTTAGGCAATTTGGTGAGAGTA 2168  
 QY 274 PheTrpThrAsnLeuTyrSerLeuThrValProPheGlyGlnLysProAsnIleAspVal 293  
 DB 2167 TATTTG--CCTTATACGTCGAGTGAATGATCTTGATTTA 2132  
 QY 294 ThrAspAlaMetValAspGlnAlaTyrAspAlaGlnArgIlePheLysGluAlaGluLys 313  
 DB 2131 ATTAACAGAGATT-----GAAACAGAGAGAAA 2108  
 QY 314 PhePheValSerValGlyLeuProAsnMet----- 323  
 DB 2107 ATTCCATTTATCTATCGGTGTTGATATGATTTGGCATATTAACATGACATTCACAGCTCGC 2048

QY 324 -----ThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGluAsn 338  
 DB 2047 TTGGAAGTGGCTGGACAGATATCTTGGAATAGAACTTAATACCAATTCACAGACAG 1988  
 QY 339 ValGlnLysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIle 358  
 DB 1987 TTTATGCAACAACTT--CATGATAGAAAGCTAAACTCTCC----- 1949  
 QY 359 LeuMetCysThrLysValThrMetAspPheLeuThrAlaHisIleGluMetGlyHis 378  
 DB 1948 -----TTAATGTCACATCTGCTGATGGGATA 1922  
 QY 379 IleGlnTyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlu 398  
 DB 1921 CGGGCTTAGAACAACCTTAT----- 1901  
 QY 399 GlyPheHisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeu 418  
 DB 1900 -----CCTCAAAATCGCA 1889  
 QY 419 LysSerIleGlyLeu-----LeuSerPro 426  
 DB 1888 AAACGGTTGGGGTTAAATGATAGAACTAGAAACCTGATTTTTCATTTTAAATCC 1829  
 QY 427 AspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeuLysGlnAlaLeuThrIle 446  
 DB 1828 TCTTTTAGGACACCTACTTTTAAAGATGTTCAI----- 1796  
 QY 447 ValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArgTrpMetValPheLysGly 466  
 DB 1795 -----TATGAACATGAAAG-- 1781  
 QY 467 GlnIleProLysAspGlnTrpMetLysLysTrpTrp--GluMetLysArgGluIleVal 485  
 DB 1780 -----CAGGAGATGATTTTGTGGATTTGATGCTGACACAGACAGACAA 1736  
 QY 486 GlyValValGluProValProHisAspGluThrTyrCysAspProAlaSerLeuPheHis 505  
 DB 1735 GGATGCTGATGATCACTT----- 1718  
 QY 506 ValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeuTyrGlnPheGlnPhe 525  
 DB 1717 -----TGCTTTTAAACCATATGACAT 1694  
 QY 526 GlnGlnAlaLeuCysGlnAla-----AlaLysHisGluGlyPro-- 538  
 DB 1693 CAGGATAGTTGTAATAATCCAGAGGTGGTTGATTTTATCAAGATATCCAGGTCTCGT 1634  
 QY 539 LeuHisLysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeu-----PheAsn 556  
 DB 1633 AGTCACCGCTACCCCTGTGTTTCA-----GGGATACATATTTATGTTGGAAAT 1583  
 QY 557 MetLeuArgLeuGlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAla 576  
 DB 1582 TCCTTTAGAGTTT-----CAACCCATTTTACACGACAGCATCAATATCGGT-- 1535  
 QY 577 LysAsnMetAsnValArgProLeuLeuAsnTyrPheGluProLeuPheThrIlePheLys 596  
 DB 1534 -----TATAGTTGG-- 1526  
 QY 597 AspGlnAsnLysAsnSerPheValGlyTyrSerThrAspTrpSerProGlyAla----- 614  
 DB 1525 -----TGGACTCATGATATCGGTGAGACATTTGTTGGG 1493  
 QY 615 -----AspGlnSerIleLysValArg-----IleSerLeuLysSer 626  
 DB 1492 GATTATGACAGACAGCTACAACTAGATGCTACAGTTTGTGTTTATAGTCGATCTACT 1433  
 QY 627 AlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetLysLeuPheArgSer 646  
 DB 1432 CGATTTCATATGTTTGAAGAGTCTTTTATAGTAAAGACCTGTGTTTTCACAGAAACA 1373  
 QY 647 ValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIle 663

Fri Mar 14 10:00:44 2003

us-09-978-385-2.p2n.rn1

Page 24

Db 1372 ACATCTAAGATTATGAGAAATAC--CTTCGTTTGAGACATCAGATGATT 1325  
:::|||||

Search completed: March 7, 2003, 09:50:13  
Job time : 166 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 7, 2003, 09:09:04 ; Search time 163 Seconds  
(without alignments)  
3298.623 Million cell updates/sec

Title: US-09-978-385-2

Sequence: 1 MSSSSMLLSLVAVTAAGT.....ISKGNPNPQNTDVCYNSF 805

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 478924 segs, 333959956 residues

Total number of hits satisfying chosen parameters: 957848

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ .p2n.model -DEV-xlp  
-Q=/cgn2\_1/USPTO.spool/US09978385/runat\_28022003.104718.1445/app\_query.fasta.1.967  
-DB-published.Applications.NA -QEMT-fastlap -SUFFIX-p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS-bits -START=1 -END=-1 -MATRIX-blosum62  
-TRANS-human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE-pct -THR MAX=100  
-THR MIN=0 -ALIGN=20 -MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MTLEN=0  
-MAXLEN=200000000 -USER=US09978385 -GCCN\_1.1.80 -runat\_28022003.104718.1445  
-NCPU=6 -ICPU=3 -NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications.NA:\*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/1/pubpna/PT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4291	100.0	3334	9	US-09-978-385-1
2	4287	99.9	3325	9	US-10-114-893-85
3	4142	96.5	3732	9	US-10-028-072-71
4	4142	96.5	3732	9	US-10-121-049-71

#### ALIGNMENTS

RESULT 1  
US-09-978-385-1  
; Sequence 1, Application US/09978385  
; Patent No. US20020177211A1  
; GENERAL INFORMATION:  
; APPLICANT: Piddington, Christopher S.  
; APPLICANT: Petrie, Charles  
; APPLICANT: Shoemaker, Kimberly E.  
; APPLICANT: Bishop, Paul D.  
; TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME  
; FILE REFERENCE: 99-24C1  
; CURRENT APPLICATION NUMBER: US/09/978,385  
; PRIOR APPLICATION NUMBER: 60/135,952  
; PRIOR FILING DATE: 2001-10-16  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: 60/151,181  
; PRIOR FILING DATE: 1999-08-27  
; PRIOR APPLICATION NUMBER: 09/563,516  
; PRIOR FILING DATE: 2000-05-03  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 3334  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (35)...(2449)



US-09-978-385-1

### Alignment Scores:

Pred. No.:	0
Score:	4291.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query match:	100.00%

```
Length: 3334
Matches: 805
```

US-09-978-385-2 (1-805) x US-09-978-385-1 (1-3334)

QY	1	MeSerSerSerSerTriPLeuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr	20
Db	35	ATGTCAAGCTCTTCTGGCTCTTCTCTGAGCTCTTGCTGTACTGCTGTGACGCCAC	94
QY	21	ILeGIuGIuGlnAlaValThrPheLeuAspLysPheAsnHISGLAlaGlnAspLeuPhe	40
Db	95	ATTGAGGAACACGGCCAAACACTTTTGTGGACAAGTTTAACCCAGGAACCGGAACTGTTC	154
QY	41	TYRGIuSerSerLeuAlaSerTriPAsnTYrAsnThrAsnILEthrGIuGIuAsnValGln	60
Db	155	TATCAAAAGTTCATCTTGCTCTTGGAATTAATACCCAAATTAATTCGAAGAAATGTCCAA	214
QY	81	GlnMeTYrProLeuGIuGlnILeGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu	100
Db	275	CAAAATGATACCACTACGAAGAAATTCACAATCTCAGTCACGTCACACTTCAGCTGACGGCTT	334
QY	101	GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrILEu	120
Db	335	CAGCAAAATGGGTCTTCAGTCTCTCTCAGAGACAAAGCAAGCAAGGGTTGACACAATCTCA	394
QY	121	AsnThrMeSerThrILEtyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu	140
Db	395	AATPACAATGACACCATCTACAGTACAGTACGAAAAAGTTTGTAACCCGAAATATCCACAAAGAA	454
QY	141	CysLeuLeuLeuGlnProGlyLeuAsnGlnILEMeAlaAsnSerLeuAspTYrAsnGlu	160
Db	455	TGCTTATTACTTGAACCAAGGTTTGAATGAATATATGCGCAAAAGTTTGAGCTTACAAATGAG	514
QY	161	ArgLeuTriPalATrPGluSerTriPArgSerGluValAlGlyLysGlnLeuArgProLeuTYr	180
Db	515	AGGCTCTGAGCGCTGGGAAAGCTGAGATGTGAGGTGCGCAACGACGCTGAGGCATTTAT	574
QY	181	GIuGIuTYrValValLeuLysAsnGlnMeAlaAlaArgAlaAsnHISTYrGluAspTYrGly	200
Db	575	GAGAGATATGTGGCTTGTGAAAAATGAAGATGGCAAGCAACATTAATGAGAGCTATAGGG	634
QY	201	AspTYrTriPArgGlyAspTYrGluValAlaAsnGlyValAlaAsnGlyTYrAspTYrSerArgLy	220
Db	635	GATTATTGGAAGGAGACTATGAAGTAAATGGGTATGATGCTTATGACTTACAGCGCGGCG	654
QY	221	GlnLeuILEglnAspValGlnHISThrPheGlnGlnILEtyrProLeuTYrGlnHISLeu	240
Db	695	CAGTTGATTTGAAGATGTGGAAACATACCTTTGAAGAGATTAACCACTTTATGAAACATCTT	754
QY	241	HISAlaTYrValAlaArgAlaLysLeuMeAlaAsnAlaTYrProSerTYrILESerProILEgly	260
Db	755	CATGGCTATGTGAGGGCAAAAGTTGATGAAGGCTTATCCCTCTATATATGCAACCAATTTGA	814
QY	261	CysLeuProAlaHISLeuLeuGlyAspMeTriPGlyArgPheThrThrAsnLeuTYrSer	280
Db	815	TGCTCTCCCTGCTCATTTGCTTGTGTGATATGTGGGAGATTTTGTGACAAACTGTACTCT	874
QY	281	LeuTriValAlaProPheGlyGlnLysProAsnILEAspValThrAspAlaMeValAspGln	300
Db	875	TTGACAGTTCCTCTTGGACAGAAACCAACAACTAAGATGTTACTGATGCAATGCTGTGACACG	934
QY	301	AlaTriPAspAlaGlnArgILEpheLysGlnAlaGlnLysPhePheValSerValGlyLeu	320

Db	935	GCCTGGGATTCGACAGAGAAATTTTCAGGAGGCCGAGAAAGTTCTTGTATCTGTGTGCTT	994
Qy	321	Probsmethylrhinglinglypethrpgluansermetleuthrasproglyasnvalgin	340
Db	995	CCTAATATGACTCAAGAGTCTGTGGAAATTTTCATGCTTAACGAGACCAGAAATGTTCAG	105
Qy	341	LysAlaValCysHisprothrAlaTTPAspleugilylGlyAspPheargileumet	360
Db	1055	AAAGAGCTCCCATCCACAGCTGTGGAGCTGGGGAAGGGCACTTCAGATCCCTTAATG	1114
Qy	361	CysTrpIysValTrpMetAspPheleuthAlaHisHISglumetClystIleGln	380
Db	1115	TGCACAAAGGTGACAAATGGAGACGCTTCTACAGCTCATATGATGGGGGATATCCAG	1174
Qy	381	TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluIlePhe	400
Db	1175	TATGATATGCGATATGCTGCGACCAACCTTTCTCTTAAGAAATGGAGCTATATAAGATTTC	1233
Qy	401	HisGluAlaValIglyIuIlemetSerLeuSerAlaAlaThrProIysHisLeuLysSer	420
Db	1235	CATGAGCTGTGGGGAAATCATGTCACTTTCGCGACCAACCTTAAGCATTTAAATACC	1294
Qy	421	IleGlyLeuLeuSerProAspPheGlnGluAspGluThrGluIleAsnPhelenu	440
Db	1295	ATTGGTCTCTGTACCCCGATTTCAGAGAACAAATGAACAGAAATTAACCTTCCTGCTC	1354
Qy	441	LysGlnAlaLeuThrIleValIglyThrLeuProPheThrTyrMetLeuGluYstrParg	460
Db	1355	AAACAGACACTCAGAGATTGTGGAGCTCTGCCATTACTTACTGATGTAGAGAAGTGGAGG	1414
Qy	461	TrpMetValPheLysGlyIuIleProLysAspGlnTrpMetLysLysTrpPgiIumet	480
Db	1415	TGGATGGCTTTAAAGGGGAAATTCCTCAAGACCAATGATGAATAAAGTGGGAGATGG	1474
Qy	481	LysArgGluIleValIglyValValGluProValProHisAspGluThrTyrCysAspPro	500
Db	1475	AAGGAGGAGATGTGGGGGTGGTGGAAACCTGTCCTCCATGATGAACATACGTACCC	1534
Qy	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgIleu	520
Db	1535	GCATCTCTGTTCCATGTTTCTTAATGATTACTCATTTATTCGATATTACACAGAGCCCTT	1594
Qy	521	TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaIleGluHisGluIleProLeuHis	540
Db	1595	TACCAATTCGAGTTTCAAGAAGCATTTGTGCAMGACGTAAACATGAAGGCCCTCTGCAC	1654
Qy	541	LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu	560
Db	1655	AAATGTACATCTCAAACTCTACAGAGCTGGGACAAAACTGTTCAATATGCTGAGGCTT	1714
Qy	561	GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetLsn	580
Db	1715	GGAAATATCAGAACCTTGAGCCCTAGCATTTGGAAATGTTGTGGAGCAAGAACCTTGAT	1774
Qy	581	ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys	600
Db	1775	GTAAGGCACAGCTCAACACTCTTTGAGCCCTTATTACTGTGCTGAAGACCAAGACAG	1834
Qy	601	AsnSerPheValIglyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal	620
Db	1835	AATATCTTTTGGGATGAGACTACCGACTGGAGTCCATATGACGACCAAGCATCAAACTG	1894
Qy	621	ArgIleSerLeuLysSerAlaLeuGlyAspIysAlaTyrGluTrpAsnAspAsnGluMet	640
Db	1895	AGGATTAAGCCCTAAATATAGCTCTTGGAGATTAAGCATTAATGAATGAACGACCAATTAATG	1954
Qy	641	TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn	660
Db	1955	TACCTGTTCGATCATCTGTGTGATATGCTATNGAGGCACTACTTTTAAAGTAAATAT	2014
Qy	661	GlnMetIleLeuPheGlyGluIleAspValArgValAlaAsnLeuLysProArgIleSer	680

Db 2015 CAGATGATCTTTTGGGAGGAGATGCGAGTGGCTAATTGAAACCAAGATCTCC 2074  
 QY 681 PheAnpPheValThraLapProLysAsnValSerAspIleIleProArgThrGluVal 700  
 Db 2075 TTTAAATTTCTTGTCTGCTACCTCACTTAAAGATGTCTGATATCATCTTCAAGACTGAAGCTT 2134  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 Db 2135 GAAAGGCCATCAGGATGTCCCGGAGCCGATCATGATGCTTCCGCTGTAATGACAAAC 2194  
 QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 Db 2195 ACCCTAGAGTTTCTGGGATACAGCCAAACACTTGGACCTCTTAACCAACCCCTGTTTCC 2254  
 QY 741 IleTrpLeuIleValPheGlyValIleValMetGlyValIleValValGlyIleValIleLeu 760  
 Db 2255 AATGTCGATTTGTTTGGAGTTGTGATGGAGTGAATGGTGGGATTCATCTCTG 2314  
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780  
 Db 2315 ATCTCAGCTGGGATCAGAGATCGGAAGAAATAAAGCAAGAGTGCAGAAATCTT 2374  
 QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800  
 Db 2375 TATGCTCTCATCGATATTAGCAAGAGAAATAATATCCAGGATTCCAAAACACTGATGAT 2434  
 QY 801 ValGlnThrSerPhe 805  
 Db 2435 GTTCAGACCTCTCTT 2449

## RESULT 2

US-10-114-893-85  
 : Sequence 85, Application US/10114893  
 : Publication No. US20020193567A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Jacobs, Kenneth  
 : APPLICANT: McCoy, John M.  
 : APPLICANT: Lavallee, Edward R.  
 : APPLICANT: Collins-Racie, Lisa A.  
 : APPLICANT: Evans, Cheryl  
 : APPLICANT: Merberg, David  
 : APPLICANT: Treacy, Maurice  
 : APPLICANT: Bowman, Michael R.  
 : APPLICANT: Spaulding, Vikki  
 : APPLICANT: Carlin-Duckett, McKeough  
 : APPLICANT: Kelleher, Kerry S.  
 : APPLICANT: Genetics Institute, Inc.  
 : TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES ENCODING THEM  
 : FILE REFERENCE: GI 6000-10A  
 : CURRENT APPLICATION NUMBER: US/10/114,893  
 : EARLIER FILING DATE: 2002-04-02  
 : EARLIER APPLICATION NUMBER: 09/413,232  
 : NUMBER OF SEQ ID NOS: 321  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 85  
 : LENGTH: 3325  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : US-10-114-893-85

## Alignment Scores:

Pred. No.: 0 Length: 3325  
 Score: 4287.00 Matches: 804  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.88% Mismatches: 0  
 Query Match: 99.91% Indels: 0  
 Db: 9 Gaps: 0

US-09-978-385-2 (1-805) x US-10-114-893-85 (1-3325)

QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaIleGlnSerThr 20  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 19 ATGCAAGCTCTTCTGCTCTCTTCTCAAGCTTGTCTTAACTGCTGCTCAAGTCCACC 78  
 QY 21 IleGluGluGlnAlaLysTrpPheLeuAspLysPheAsnIleGluAlaGluAspLeuPhe 40  
 Db 79 ATTGAGAACAGGCGCAAGACATTTTGTGACAGATTATACCAACGAGGCGCAAGACGTGTTG 138  
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGln 60  
 Db 139 TATCAAGGTCACCTTCTTGTGAATTAACACCAATATTAATCTAGTAGAGGAATGTCCAA 198  
 QY 61 AsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80  
 Db 199 AACTGATATATGTGTGGGCAAAATGTCTGCTTTTAAAGAACAGTCCACACTTGCC 258  
 QY 81 GlnMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeu 100  
 Db 259 CAATGTATCCACTACACAAAGAAATTCAGAAATTCACAGCTTCAGCTTCAGCTGCGCTCT 318  
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120  
 Db 319 CAGCAAAATGGGTCTTCAAGTCTCTCAGAGAACAGCAAGAGGTTGAACACAAATTTCTA 378  
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140  
 Db 379 AATACATATGAGCACCATCTACAGTACGGAAGAGTTTGTACCCCATATATCCACAGAA 438  
 QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160  
 Db 439 TGCATTATTACTTAACACCGGTTGAATGAATATGCAACAGTTTATGACTACATATGAG 498  
 QY 161 ArgLeuTrpAlaTrpGlnSerTrpArgSerGluValGlyGlnLeuArgProLeuTyr 180  
 Db 499 AGGCTCTGGGCTTGGGAAACCTGGAGATCTGAGAGTGGCAAGCAAGCTGAGGCCATTATAT 558  
 QY 181 GluGlnTyrValValLeuLysAsnGluMetAlaArgAlaAsnIleThrGluAspTyrGly 200  
 Db 559 GAAGAGTATGTGCTTGAAGAAATGAGATGCGCAAGCAAAATCATATATGAGGACTATGGG 618  
 QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220  
 Db 619 GATTATTGGAGAGAGCATATGAGTAAGTAAGGGGTGATGCTATACATACAGCCCGGCG 678  
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240  
 Db 679 CAGTGTATTAACATGTGGAGAACATACCTTGAAGAGATTAACCATTAATATGACACTCTT 738  
 QY 241 HisAlaTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260  
 Db 739 CATGCTATGTGAGGCAAGTGTGATGAATGCCTATCTCTATATACAGTCCAAATTGGA 798  
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheThrAsnLeuTyrSer 280  
 Db 799 TGCCTCCCTGCTCATTTGCTGTGTGATATGTGGGTGATTTGGCAAAATCTGTACTCT 858  
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 Db 859 TTGACAGTTCCTTGTGAGCAAGAAACCAACATATGATCTTACTGATCATGTGTGACCG 918  
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320  
 Db 919 GCCTGGATGCACAGAAATATTCAGAGGAGGCCGAGAAAGTTCTTGTATCTTGTGGTCTT 978  
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyLysValGln 340  
 Db 979 CTTAATATATCTCAAGGATTTCTGGGAAATTTCCATGCTTAACGAGCCAGAAATGTTCAG 1038  
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 Db 1039 AAGGCAAGTGTGACATCCACAGAGTGTGGACCTGGGGAAGGAGGACTCGATGCTCTTATG 1098  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGluMetGlyHisIleGln 380  
 Db 1099 TGCACAAGGTGACATGAGAGGACTTCTGACAGCTCATGATGATGGGATATATCCAG 1158

QY 381 TYRAspMetAlaTyrAlaIaInProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400  
 DB 1159 TATGATATGCGATATGCTCACAACCTTTTCTGCTAGAGAAAGAGCTATATAGAGATTC 1218  
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaIaIaThrProLysHisLeuLysSer 420  
 DB 1219 CATGAGAGCTGTGGGAAATCATGTCTTCTGGAGCCACACCTAAGCATTTAAATATCC 1278  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440  
 DB 1279 ATTTGCTCTCTGTCACCCATTTTCAAGAAACAAAGAAATTAACCTTCTGCTCC 1338  
 QY 441 LysGluAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluThrTrpArg 460  
 DB 1339 AAACAAGCACTACAGATTTGTGGAGCTTGCCATTACTTACATGTAGAGAAAGTGGAGG 1398  
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpTrpGluMet 480  
 DB 1399 TGGATGCTCTTAAAGGGGAAATTTCCCAAGAACACAGTGGATGAAAAAGTGTGGAGATG 1458  
 QY 481 LysArgGluIleValGlyValGluProValProHisAspGluThrTyrCysAspPro 500  
 DB 1459 AACCGAGAGATAGTGGGTGGTGGAACTGTGCCCATGATGAACACATCTGTACCC 1518  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520  
 DB 1519 GCATCTCTGTCATTTCTTAATGATTACTCATTCATTCGATATTACCAAGACCTT 1578  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuGlyGlnAlaIaIaLysHisGluIleProLeuHis 540  
 DB 1579 TACCAATTCAGATTCAAGAACACTTTGTCAAGCAGCTAAACATGAAGCCCTCTGAC 1638  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560  
 DB 1639 AAATGTGACATCTCAACACTCTACAGAACTGGACAGAACTGTCAATATGCTGAGGCTT 1698  
 QY 561 GlyLysSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetLsn 580  
 DB 1699 GGAATATCAGAACCTCGAGACCTGAGCATTTGGAATAATGTTGTAGGAGCAAGAACATGAT 1758  
 QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTripleLysAspGlnLys 600  
 DB 1759 GTTAGGCCACTGCTCAACACTTTGAGCCCTTATTACTGCTGTAACCAAGCAAGCAAG 1818  
 QY 601 AsnSerPheValGlyTyrPheSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620  
 DB 1819 AATTCTTTTGTGGATGGAGTACCGACTGAGTCCATATGCGAGACCAAGCATCAAGTG 1878  
 QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluThrPasnAspAsnGluMet 640  
 DB 1879 AGGATAGCCCTAAATACAGCTTGGAGATTAAGCATATGATGAAGCAACATGAATG 1938  
 QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660  
 DB 1939 TACCTTTCGATCATCTGTGCATATGATGAGGCACTACTTTTAAAGTAAAAAAT 1998  
 QY 661 GluMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680  
 DB 1999 CAGATGATTTCTTTTGGGAGAGAGATGCGAGTGGCTAATTTGAAACAGAAATCTCC 2058  
 QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700  
 DB 2059 TTTAATTTCTTTGCTACCTGACCTAAATGTCGTGATATCATCTTCTGAACTGAAGTT 2118  
 QY 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720  
 DB 2119 GAAAGGCGCATTCAGAGATGCCGAGCCGATCAATGATGCTTCCGTGTGAATGCAAC 2178  
 QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740  
 DB 2179 AGCTTAGAGTTCTTGGGAGTACAGCAACACTTGGAGCTCTTCAACACAGGCCCTCTGTTCC 2238

QY 741 IleTrpLeuIleValPheGlyValAlaMetGlyValIleValValGlyIleValIleLeu 760  
 DB 2239 ATATGGCGGATGTTGTTTGGAGCTGTGATGGAGATGATATGATGTCATTTGCATCCTG 2298  
 QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnPro 780  
 DB 2299 ATCTTCACTGGGATCAGAGATCGAGAGAGAAAAATTAACCAAGAACTGGAGAAAAATCCT 2358  
 QY 781 TyrAlaSerIleAspIleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAsp 800  
 DB 2359 TATGCTTCATCATGATTATTAGCAAGAGAAATATATCCAGATTTCCAAACACTGATGAT 2418  
 QY 801 ValGlnThrSerPhe 805  
 DB 2419 GTTCAGACCTCCTT 2433  
 RESULT 3  
 US-10-028-072-71  
 : Sequence 71, Application US/10028072  
 : Publication No. US20030004311A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Baker, Kevin P.  
 : APPLICANT: Beresini, Maureen  
 : APPLICANT: Deforge, Laura  
 : APPLICANT: Desnoyers, Luc  
 : APPLICANT: Filvaroff, Ellen  
 : APPLICANT: Gao, Wei-Qiang  
 : APPLICANT: Gerritsen, Mary E.  
 : APPLICANT: Goddard, Audrey  
 : APPLICANT: Godowski, Paul J.  
 : APPLICANT: Gurney, Austin L.  
 : APPLICANT: Sherwood, Steven  
 : APPLICANT: Smith, Victoria  
 : APPLICANT: Stewart, Timothy A.  
 : APPLICANT: Tumas, Daniel  
 : APPLICANT: Watanabe, Colin K  
 : APPLICANT: Wood, William  
 : APPLICANT: Zhang  
 : TITLE OF INVENTION:  
 : FILE REFERENCE:  
 : CURRENT APPLICATION NUMBER: US/10/028, 072  
 : CURRENT FILING DATE: 2001-12-19  
 : PRIOR APPLICATION NUMBER: 60/049911  
 : PRIOR FILING DATE: 1997-06-18  
 : PRIOR APPLICATION NUMBER: 60/056974  
 : PRIOR FILING DATE: 1997-08-26  
 : PRIOR APPLICATION NUMBER: 60/059113  
 : PRIOR FILING DATE: 1997-09-17  
 : PRIOR APPLICATION NUMBER: 60/059115  
 : PRIOR FILING DATE: 1997-09-17  
 : PRIOR APPLICATION NUMBER: 60/059117  
 : PRIOR FILING DATE: 1997-09-17  
 : PRIOR APPLICATION NUMBER: 60/059122  
 : PRIOR FILING DATE: 1997-09-17  
 : PRIOR APPLICATION NUMBER: 60/059184  
 : PRIOR FILING DATE: 1997-09-17  
 : PRIOR APPLICATION NUMBER: 60/059263  
 : PRIOR FILING DATE: 1997-09-18  
 : PRIOR APPLICATION NUMBER: 60/059352  
 : PRIOR FILING DATE: 1997-09-19  
 : PRIOR APPLICATION NUMBER: 60/059588  
 : PRIOR FILING DATE: 1997-09-19  
 : PRIOR APPLICATION NUMBER: 60/059836  
 : PRIOR FILING DATE: 1997-09-24  
 : PRIOR APPLICATION NUMBER: 60/062250  
 : PRIOR FILING DATE: 1997-10-17  
 : PRIOR APPLICATION NUMBER: 60/062285  
 : PRIOR FILING DATE: 1997-10-17  
 : PRIOR APPLICATION NUMBER: 60/062287  
 : PRIOR FILING DATE: 1997-10-17  
 : PRIOR APPLICATION NUMBER: 60/062814  
 : PRIOR FILING DATE: 1997-10-24  
 : PRIOR APPLICATION NUMBER: 60/062816

1	PRIOR FILING DATE: 1997-10-24
2	PRIOR APPLICATION NUMBER: 60/063045
3	PRIOR FILING DATE: 1997-10-24
4	PRIOR APPLICATION NUMBER: 60/063082
5	PRIOR FILING DATE: 1997-10-31
6	PRIOR APPLICATION NUMBER: 60/063127
7	PRIOR FILING DATE: 1997-10-24
8	PRIOR APPLICATION NUMBER: 60/063327
9	PRIOR FILING DATE: 1997-10-27
10	PRIOR APPLICATION NUMBER: 60/063329
11	PRIOR FILING DATE: 1997-10-27
12	PRIOR APPLICATION NUMBER: 60/063550
13	PRIOR FILING DATE: 1997-10-28
14	PRIOR APPLICATION NUMBER: 60/063561
15	PRIOR FILING DATE: 1997-10-28
16	PRIOR APPLICATION NUMBER: 60/063704
17	PRIOR FILING DATE: 1997-10-29
18	PRIOR APPLICATION NUMBER: 60/063733
19	PRIOR FILING DATE: 1997-10-29
20	PRIOR APPLICATION NUMBER: 60/063735
21	PRIOR FILING DATE: 1997-10-17
22	PRIOR APPLICATION NUMBER: 60/064248
23	PRIOR FILING DATE: 1997-11-03
24	PRIOR APPLICATION NUMBER: 60/064809
25	PRIOR FILING DATE: 1997-11-07
26	PRIOR APPLICATION NUMBER: 60/065186
27	PRIOR FILING DATE: 1997-11-12
28	PRIOR APPLICATION NUMBER: 60/065846
29	PRIOR FILING DATE: 1997-11-17
30	PRIOR APPLICATION NUMBER: 60/066364
31	PRIOR FILING DATE: 1997-11-21
32	PRIOR APPLICATION NUMBER: 60/066453
33	PRIOR FILING DATE: 1997-11-24
34	PRIOR APPLICATION NUMBER: 60/066511
35	PRIOR FILING DATE: 1997-11-24
36	PRIOR APPLICATION NUMBER: 60/066770
37	PRIOR FILING DATE: 1997-11-24
38	PRIOR APPLICATION NUMBER: 60/069212
39	PRIOR FILING DATE: 1997-12-11
40	PRIOR APPLICATION NUMBER: 60/069278
41	PRIOR FILING DATE: 1997-12-11
42	PRIOR APPLICATION NUMBER: 60/069334
43	PRIOR FILING DATE: 1997-12-11
44	PRIOR APPLICATION NUMBER: 60/069694
45	PRIOR FILING DATE: 1997-12-16
46	PRIOR APPLICATION NUMBER: 60/072320
47	PRIOR FILING DATE: 1998-01-23
48	PRIOR APPLICATION NUMBER: 60/073512
49	PRIOR FILING DATE: 1998-02-04
50	PRIOR APPLICATION NUMBER: 60/074086
51	PRIOR FILING DATE: 1998-02-09
52	PRIOR APPLICATION NUMBER: 60/074092
53	PRIOR FILING DATE: 1998-02-09
54	PRIOR APPLICATION NUMBER: 60/077791
55	PRIOR FILING DATE: 1998-03-12
56	PRIOR APPLICATION NUMBER: 60/078910
57	PRIOR FILING DATE: 1998-03-20
58	PRIOR APPLICATION NUMBER: 60/079294
59	PRIOR FILING DATE: 1998-03-25
60	PRIOR APPLICATION NUMBER: 60/079663
61	PRIOR FILING DATE: 1998-02-27
62	PRIOR APPLICATION NUMBER: 60/079728
63	PRIOR FILING DATE: 1998-03-27
64	PRIOR APPLICATION NUMBER: 60/080165
65	PRIOR FILING DATE: 1998-03-31
66	PRIOR APPLICATION NUMBER: 60/081203
67	PRIOR FILING DATE: 1998-04-09
68	PRIOR APPLICATION NUMBER: 60/081229
69	PRIOR FILING DATE: 1998-04-09

1	PRIOR APPLICATION NUMBER: 60/081655
2	PRIOR FILING DATE: 1998-04-14
3	PRIOR APPLICATION NUMBER: 60/081817
4	PRIOR FILING DATE: 1998-04-15
5	PRIOR APPLICATION NUMBER: 60/081818
6	PRIOR FILING DATE: 1998-04-15
7	PRIOR APPLICATION NUMBER: 60/082999
8	PRIOR FILING DATE: 1998-04-24
9	PRIOR APPLICATION NUMBER: 60/083322
10	PRIOR FILING DATE: 1998-04-28
11	PRIOR APPLICATION NUMBER: 60/083545
12	PRIOR FILING DATE: 1998-04-29
13	PRIOR APPLICATION NUMBER: 60/084600
14	PRIOR FILING DATE: 1998-05-07
15	PRIOR APPLICATION NUMBER: 60/084627
16	PRIOR FILING DATE: 1998-05-07
17	PRIOR APPLICATION NUMBER: 60/084637
18	PRIOR FILING DATE: 1998-05-07
19	PRIOR APPLICATION NUMBER: 60/085149
20	PRIOR FILING DATE: 1998-05-12
21	PRIOR APPLICATION NUMBER: 60/085323
22	PRIOR FILING DATE: 1998-05-13
23	PRIOR APPLICATION NUMBER: 60/085338
24	PRIOR FILING DATE: 1998-05-13
25	PRIOR APPLICATION NUMBER: 60/085339
26	PRIOR FILING DATE: 1998-05-13
27	PRIOR APPLICATION NUMBER: 60/085579
28	PRIOR FILING DATE: 1998-05-15
29	PRIOR APPLICATION NUMBER: 60/085697
30	PRIOR FILING DATE: 1998-05-15
31	PRIOR APPLICATION NUMBER: 60/085704
32	PRIOR FILING DATE: 1998-05-15
33	PRIOR APPLICATION NUMBER: 60/086414
34	PRIOR FILING DATE: 1998-05-22
35	PRIOR APPLICATION NUMBER: 60/086430
36	PRIOR FILING DATE: 1998-05-22
37	PRIOR APPLICATION NUMBER: 60/087106
38	PRIOR FILING DATE: 1998-05-28
39	PRIOR APPLICATION NUMBER: 60/088026
40	PRIOR FILING DATE: 1998-05-04
41	PRIOR APPLICATION NUMBER: 60/088730
42	PRIOR FILING DATE: 1998-06-10
43	PRIOR APPLICATION NUMBER: 60/088741
44	PRIOR FILING DATE: 1998-06-10
45	PRIOR APPLICATION NUMBER: 60/088810
46	PRIOR FILING DATE: 1998-06-10
47	PRIOR APPLICATION NUMBER: 60/088858
48	PRIOR FILING DATE: 19/98-06-11
49	PRIOR APPLICATION NUMBER: 60/089532
50	PRIOR FILING DATE: 1998-06-17
51	PRIOR APPLICATION NUMBER: 60/089599
52	PRIOR FILING DATE: 1998-06-17
53	PRIOR APPLICATION NUMBER: 60/089907
54	PRIOR FILING DATE: 1998-06-18
55	PRIOR APPLICATION NUMBER: 60/089947
56	PRIOR FILING DATE: 1998-06-19
57	PRIOR APPLICATION NUMBER: 60/090349
58	PRIOR FILING DATE: 1998-06-23
59	PRIOR APPLICATION NUMBER: 60/090429
60	PRIOR FILING DATE: 1998-06-24
61	PRIOR APPLICATION NUMBER: 60/090445
62	PRIOR FILING DATE: 1998-06-24
63	PRIOR APPLICATION NUMBER: 60/090538
64	PRIOR FILING DATE: 1998-06-24
65	PRIOR APPLICATION NUMBER: 60/090863
66	PRIOR FILING DATE: 1998-06-26
67	PRIOR APPLICATION NUMBER: 60/091360
68	PRIOR FILING DATE: 1998-07-01
69	PRIOR APPLICATION NUMBER: 60/091519
70	PRIOR FILING DATE: 1998-07-02
71	PRIOR APPLICATION NUMBER: 60/091982
72	PRIOR FILING DATE: 1998-07-07

Alignment Scores:

Pred. No.:	0	Length:	373
Score:	4142.00	Matches:	802
Percent Similarity:	85.33%	Conservative:	1
Best local Similarity:	85.23%	Mismatches:	2
Query Match:	96.53%	Indels:	138
DB:	9	Gaps:	1

US-09-978-385-2 (1-805) x US-10-028-072-71 (1-3732)

Oy	1	MeserSerSerSerTrpLeuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr	20
Db	40	ATGTCAACCTCTTCTCGCTCTCTTCTAGCCTTGTTCTGTAACTGCTGCTACGCC	99
Oy	21	IleGluGlnAlaAlaLysThrPheLeuAspLysPheAsnHisGluAlaGlnAspLeuPhe	40
Db	100	ATTGGAGCAACAGGCCAAGACATTTTGTGACAAAGTTTAACCAAGCAAGCCGAAGCACTGTTTC	159
Oy	41	TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnLleThrGluGlnValGln	60
Db	160	TATCAAAATTCACCTGCTCTTGGAAATTAATACCAACCAATATTCGTGAAGGAATGTCCAA	219
Oy	61	AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla	80
Db	220	AACATGATTAATGCTGGGGCAAAATGGCTGCCCTTTTAAAGAACAGTCCACACTTGCC	279
Oy	81	GlnMetTyrProLeuGlnLleGlnAsnLeuThrValLysLeuGlnLeuAlaLeu	100
Db	280	CAAACTATCCACACTACAGAAATTCAGAAATCTCAGCTCAAGGCTTCAGCTGCAGGCTCTT	339
Oy	101	GlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrLleLeu	120
Db	340	CAGCAAAATGGGCTCTCAGTCTCTCGAAGAACAGCAAGCAACGGTTGAACACAAATCTA	399
Oy	121	AsnThrMetSerThrLleTyrSerThrGlyValLysAsnProAspAsnProGlnGlu	140
Db	400	AATACAAAGACACCATTTACAGTACGTACGAAAAATTTGTAAACCCAGATTAATCCACAAGAA	459
Oy	141	CysLeuLeuLeuGluProGlyLeuAsnGlnLleMetAlaAsnSerLeuAspTyrAsnGlu	160
Db	460	TGCTTATTAATCTTGAACCGAGTTTGAATGAATATATGCAACAAAGTTTAAGACTCAAAAGAG	519
Oy	161	ArgLeuTrpAlaTrpGlnLysSerTrpArgSerGluValGlyGlnLeuArgProLeuTyr	180
Db	520	AAGCCTCTGGGCTGGGAAAGCTGGAGATCTAGGTGGCAAGCAGCTGAGGCCATTATAT	579
Oy	181	GlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGluAspTyrGly	200
Db	580	GAAAGATGTGTGCTTTAAAAAATGAGATGGCAAGAGCAAAATCTATTGAGGACTATGGG	639
Oy	201	AspTyrTrpArgGlyAspTyrGluValAsnGlnValAspGlyTyrAspTyrSerArgGly	220
Db	640	GATTTATTTGGAGAGAGACTTAAGAAGTAATATGGGCTAGATGGCTATGACTACAGCCGCGGC	699
Oy	221	GlnLeuLleGluAspValGlnHisThrPheGlnGlnLleLysProLeuTyrGlnHisLeu	240
Db	700	CAGTTGATTTGAAGAGTGAACATPACCTTTAAAGAGATTAACCACTTATATAGAACATCTT	759
Oy	241	HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrLleSerProLleGly	260
Db	760	CATGCCCTATGTAGAGGCAAAAGTGTATGATGATGCCCTTACTCTCTTATACAGGCCAAATTGGA	819
Oy	261	CysLeuProAlaHisLeuLeuGlnLysMetTrpGlyArgPheTrpThrAsnLeuTyrSer	280
Db	820	TGCCCTCCGCTCANTTCTGCTGTGATATGTGGGGTAGATTTTGGACAAATCTGACTCT	879
Oy	281	LeuThrValProPheGlyGlnLysProAsnLleAspValThrAspAlaMetValAspGln	300
Db	880	TTTGACAGTTCCTCTTGGACAGAAACCAAAACATATGTTACTGTATGCAATGTGTGACACAG	939
Oy	301	AlaTrpAspAlaGlnArgLlePheLysGlnLalGlnLysPhePheValSerValGlyLeu	320
Db	940	GCTTGGGATGCAACAAGATATTTCAAGAGGCCGAGAAATGTTCTTGTATCTGTGTGGCTT	999

OY	321	ProhomethrhcnglnlyPheRtpgluaNserMeLleuThAspProGlyAsnValGln	340
Db	1000	CCNATATGTGACTGAAGGATTTCTGGGAAATTTCCATGCTTACAGGACCAGAAATGTTTCAG	1059
OY	341	LysAlaValCysHisPheThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet	360
Db	1060	AAAGACGTGTCACATGCCACAGCTTGGAGACCTGGGGAAAGGGCGACTTCGAGATCTTATG	1119
OY	361	CysThrIysValThrMetAspAspPheLeuThrAlaHisIleGluMetGlyHisIleGln	380
Db	1120	TGCAAAAGGTGCAATGAGCGACTTCGACAGCTTCATCATGAAATGGGCGCATATCCAG	1179
OY	381	TyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe	400
Db	1180	TATGATATGGCATATGCTGTCACAACTTTTCTGTAAAGAAATGGAGTATGAAGATTC	1239
OY	401	HisGluAlaValGlyGluIleMetSerLeuSerAlaIleThrProLysHisLeuLysSer	420
Db	1240	CATTAACCTGTTGGGGAAATCAATGATCACTTCTGACGCACACCTAACCATTTTAAATCC	1299
OY	421	IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu	440
Db	1300	ATTGCTTTCTGTACCCCATTTTCAAGAAGCAATGAAGCAATAACTTCCTGCTC	1359
OY	441	LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg	460
Db	1360	AAACAAACACTGCAGATTTGTTGGGACTCTGCCATTACTACATGTTAGAGAAAGGAGG	1419
OY	461	TrpMetValPheLysGlyLysIleProLysAspGlnTrpMetLysLysTrpGluMet	480
Db	1420	TGGATGCTTTTAAAGGGGAAATTTCCCAAGACAGCTGATGAATAAAGTGGGGAAGTG	1479
OY	481	LysArgGluIleValGlyValValGluProValProHisAspGluThrTyrCysAspPro	500
Db	1480	AAGCGAGAGATAGTTGGGGTGGTGAACCTGTGCCCATGATGAAGAACTACTGTGACCCC	1539
OY	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu	520
Db	1540	GCATCTCTGTTCCATGTTTGTGATGATTCCTCATTCATTCGATATTCACAGAGGCCCTT	1599
OY	521	TyrGlnPheGlnPheGlnGluAlaLeuCysGlnIleAlaIleValHisGluGlyProLeuHis	540
Db	1600	TACCAATTTCCAGTTTCAAGAAGCACTTGTCAAGCAGCTAACATGAAGGCCCTGTGCAC	1659
OY	541	LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe-----	555
Db	1660	AAATGTACATCTCAAACTCTACAGAAGCTGACAGAAACTGTT- GTAAGAAATACCTCA	1718
OY	555	-----	555
Db	1719	AAATGTTGAACCTCTCCTAGTATTCAGTATTACTCATTTCCATGGCTTAGGTTTGATTTG	1778
OY	555	-----	555
Db	1779	ATTCTTGTGCTCAAAAAGAAATTTTATAGCCCTCAAAATGCTCATTTACAAACCAAA	1838
OY	555	-----	555
Db	1839	CATTAAATTGTGTGTCAGACAGGAACTGACCATCAACAAATTGGGTGGCCACTCTTT	1898
OY	555	-----	555
Db	1899	TTTCGCCATCATACACAGCCCTCTCTTCCTCGTGAATTTGGAAAGAAAGACGGTTTAG	1958
OY	555	-----	555
Db	1959	GATGGAATATATCTGTTAATATGCAATCTTTCTTATCTGCGAAGAAATTTAGCCAA	2018
OY	555	-----	555
Db	2019	GTCAAAGAGAAACCATGATCATCATGATGTAAATATATGTACATCTGGAAACCCCTTAA	2078



```

|||||
Db 700 CAGTTGATGAGATGATGACATACCTTTGAAGAGATTAAACATATATGACATCTT 759
QY 241 HIsAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 760 CATGCTTAATGTGGGCAAGATTGATGATGCTTACCTTCTTATATCATGCTCAATTTGA 819
QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTyrGlyArgPheTyrPheTrpHisLeuTyrSer 280
Db 820 TGGCTCCCTGCTGCATTTGCTTGTATGTGGTAGAGTTTGTGGCAAAATGTGACTCT 879
QY 281 LeuThrValProPheGlyLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 880 TTGACAGTTCCCTTGTGACAGAAACCAACATAGATGTACTGATGATCAATGCTGAGCCAG 939
QY 301 AlaTyrAspAlaGlnArgIlePheLysGluAlaGlnLysPhePheValSerAlaGlyLeu 320
Db 940 GCTTGGATGACACAGAAATATTCAAGAGCCGAGAAAGTTCTTGTATCTGTGCTCTT 999
QY 321 ProAsnMetThrGlnGlyPheTyrGlyAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 1000 CCTAATATGACTCAAGGATTTGGGAAATTTCCATGCTTACGAGCCAGCAAAATGTTGAG 1059
QY 341 LysAlaValCysHisProThrAlaTyrAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1060 AAGACGCTGCTGCATCCACAGCTTGGGACCTGGGGAGGCGGACTTCAGATCCTTATG 1119
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380
Db 1120 TGCACAAAGGTGACAAATGACACATCTTCTGACAGCTCATAGAAATGGGGATATTCAG 1179
QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuAsnArgAsnGlyAlaAsnGlnLysPhe 400
Db 1180 TATGATATGCTATGCTCACAACCTTTCTGCTAAGAAATGAGCTAATGAGGATTC 1239
QY 401 HisGluAlaValGlyGlnIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
Db 1240 CATGAAAGCTGTGGGAAATCATGCTTCTTGCAGCCACACCTTCAACATTTAAATCC 1299
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGlnThrGlnIleAsnPheLeuLeu 440
Db 1300 ATTGCTCTCTGTACCCCATTTTCAAGACACATGAAACGAATTAACCTCTGCTC 1359
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTyrArg 460
Db 1360 AAACAAGACACTCAGATTTGTTGGACTCTGCCATTTACTTACATGTAGAGAGTGGAG 1419
QY 461 TrpMetValPheLysGlyGlnIleProLysAspGlnTrpMetLysLysTyrTrpGlnMet 480
Db 1420 TGGATGCTCTTTAAAGGGGAAATTTCCAAAGACAGTGGATGAAAAGTGGTGGGAGT 1479
QY 481 LysArgGlnIleValGlyValAlaGlnProValProHisAspGlnThrTyrCysAspPro 500
Db 1480 AAGCAGAGATATGTTGGGGTGGTGGACCTGTGCCCATGATGAACATACGTGAGCCCC 1539
QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
Db 1540 GCATCTCTGTTCCATGTTTGCATGATTAATCATTCATTCATTCATTCATTCATTCATTC 1599
QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHis 540
Db 1600 TACCAATTCACATTTCAAGAGACACTTGTCAAGCAGCTTAAACATGAAAGGCCCTG 1659
QY 541 LysCysAspIleSerAsnSerThrGlnAlaGlnLysLeuPhe----- 555
Db 1660 AAATGTGACATCTCAAACTTACAGAACTGAGAGAACTGTT- GTAAGAAATACCTCA 1718
QY 555 ----- 555
Db 1719 AAATGTGACCTCTCTAGATTACATTAATTCATTCATTCATTCATTCATTCATTCAT 1778
QY 555 ----- 555

|||||
Db 1779 ATTCTTGTCTTAAGAAAGAAATTTTATGGCTCAAAATGTCTCATTTACAACCAA 1838
QY 555 ----- 555
Db 1839 CATTTAATTTGTGTGACAGACAGAACTAGACCATACACAAATTTGGTGGCCACTCTT 1898
QY 555 ----- 555
Db 1899 TTTCCTCATATACATACAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1958
QY 555 ----- 555
Db 1959 GGTGAAATATATCTGTTAATATGATCTTTTCTTATCTGCGAAGCAAAATTTAGCCA 2018
QY 555 ----- 555
Db 2019 GTCAAGAGAGAAACCATATAGATCATATATATATATATATATATATATATATATAT 2078
QY 556 ----- 556
Db 2079 AAGGCCTGAAACCCCTTTTGTGTAGCAATATGCTGAGGCTTGGAAATATGAAACCC 2138
QY 566 TrpThrLeuAlaLeuGluAsnValAlaGlyAlaLysAsnMetAsnValArgProLeuLeu 585
Db 2139 TGGACCTAGCATGGAAATGTTGTAGAGC- AAGAAATGAAATGTAAGGCCACTGCTC 2197
QY 586 AspTyrPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGly 605
Db 2198 AACTACTTGTGACCTTATTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2257
QY 606 TrpSerThrAspTyrSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeuLys 625
Db 2258 TGGAGTACCGCATGATCCATATGACAGCCAAAGATCAAAAGTAGAGATAGCCTAAT 2317
QY 625 SSerAlaLeuGlyAspLysAlaTyrGlnTyrPheAsnAspGlnMetTyrLeuPheArgSe 645
Db 2318 ATCAGCTCTTGGAGATTAACCATATGATGAAAGCAGCAAAATGATGATGATGATGATG 2377
QY 645 rSerValAlaTyrAlaMetArgIleThrPheLeuLysValLysAsnGlnMetIleLeuPhe 665
Db 2378 ATCTGTTGCATATGCTATGAGGAGTACTTTTAAAGTAAATAACATGATGATGATGATG 2437
QY 665 eGlyGlnGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheVal 685
Db 2438 TGGGAGAGAGATGTGGAGTGGCTAATTTGAACCAAGATCTCCTTTAATTTCTTTGT 2497
QY 685 IThrAlaProLysAsnValSerAspIleIleProArgThrGluValGlnLysAlaIleArg 705
Db 2498 CACTGCACCTTAAAAATGTCTGATATCATCTTCAAGACGAGCTGAAAGGCCATCAG 2557
QY 705 gMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGlnPheLe 725
Db 2558 GATGTCCCGGAGCCGTATCATATGATGCTTCCGTGGAAGACAAAGCTGAGATTTCT 2617
QY 725 uGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTrpLeuIleVal 745
Db 2618 GGGGATACACCAACAACTTTGGACCTCTTAACAGCCGCCCTGTTCCATATGCTGATTT 2677
QY 745 LpneGlyValAlaMetGlyValIleValAlaGlyIleValIleLeuIlePheThrGlyIle 765
Db 2678 TTTTGGAGTTGTGATGGAGTGTAGTGTGATGTTGATGATGATGATGATGATGATGAT 2737
QY 765 eArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnProTyrAlaSerIleAs 785
Db 2738 CAGAGATCGGAGAGAGAAATTAAGCAAGAACTGAGAGAAATCTTAAGCCTCATCA 2797
QY 785 PileSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPhe 805
Db 2798 TATTAGCAAGAGAGAAATATATCAGAGATTCCAAAACACTGATGATGATGATGATGAT 2857
QY 805 e 805
Db 2858 T 2858

```





Qy	501	AlaserleupPheHsValSerAsnAspTyrSerPheIleArgTyrThrThrArgHleu	520
Db	1540	GCATCTCGATCCATGTTTCTTGATGATTAATCTCAATTCGATATTACAAAGACCTT	1595
Qy	521	TyrGlnPheGlnPheGlnGlnValLeuCySGlnAlaAlaIleHisGlnLeuProleuHis	540
Db	1600	TACCAATTTCCAGTTTCAAGAAGACCTTTGTCCAGCAGCATTAACATGAAGCCCTCTGCAC	1655
Qy	541	LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe-----	555
Db	1660	AAAGTGACATCTCAAACTCTACAGAAGCTGGACAGAAACTGTT--GTAGAATAACTTCA	1715
Qy	555	-----	555
Db	1719	AAATGTTGAACCTCTCCTAGTATTCAGTATTACTCATTTTCCATGCTAGGTTTGATTTTG	1775
Qy	555	-----	555
Db	1779	ATTTCCTTTGTTTAAAAAGAAATTTTATGGCCCTCAAAATGTCTTCATTTCACAAACCAA	1835
Qy	555	-----	555
Db	1839	CATTTAATTTGTGGTCAGACAGGAACTAGACATCAACAATTTGGTGGCCACCTCTT	1895
Qy	555	-----	555
Db	1899	TTTCCTCATCATCACTACAGACCCCTCTTCTCTGTAAATTGGAAAGAAAGCGGTTTAG	1955
Qy	555	-----	555
Db	1959	GGTGGAAATATATCTGTTAAATATGATCTTTCTTATCTGCCGAAGCAAAATTTAGCCA	2015
Qy	555	-----	555
Db	2019	GTCAAGAGAAAGAAACCATGATCATGATGTAAATATATGTACATCTGGAACCCCTCAA	2075
Qy	556	-----AsmleuArgleuGlnLysSerGlnPro	565
Db	2079	AAGGCCCTGAACCCCTTTTTTGTGTGAGCAAAATCTGTGGCTGGAAATCAGAACCC	2135
Qy	566	TrpThrLeuAlaLeuGlnAsnValAlaGlyAlaLysAsnMetAsnValArgProleuLeu	585
Db	2139	TGGACCTCAGATTTGGAAATGTTTGTAGGAC--AAGACACTGATATTAAGCCACTGCTC	2195
Qy	586	AsnTyrPheGlnProleuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGly	605
Db	2198	AACTACTTTGAGCCCTTATTATTTACCTGGCTGAAGACAGAACAGAATCTTTTGGGGA	2255
Qy	606	TrpSerThrAspTrpSerProTyrAlaAsp--GlnSerIleLysValArgIleSerLeuLys	625
Db	2258	TGGAGTACCCGACTGGAGCTCATATGCGAGCCCAAAACATCAAAAGTAGGATTAAGCTTAA	2315
Qy	625	SerAlaLeuGlnLysAspLysAlaTyrGlnTrpAsnAspAsnGlnMetTyrLeuPheArgSe	645
Db	2318	ATCAGCTCTTGGAGATTAAGCATATGTAATGGAACGACAAATGAATTTACTTGTCCGATC	2375
Qy	645	rSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPhe	665
Db	2378	ATCTGTTCGATATCTCATGAGCGCGAGTCTTTTTAAAAAGTAAAAAATCAGATGATCTTTT	2435
Qy	665	eGlyGlnGlnLysAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePhe	685
Db	2438	TGGGAGAGAGATTTGGAGTGGCTAATTGGAACCAAGAAATCTCTTTAAATTTCTTTGT	2495
Qy	685	lThrAlaProLysAsnValSerAspIleIleProArgThrGlnValGlnLysAlaIleArg	705
Db	2498	CACCTGCACCTAAAAATGTCTCATATCATCTTCAACTGAACTGAAAGTTGAAAGGCCCATCG	2555
Qy	705	GmetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGlnPheLe	725
Db	2558	GATGTCGCGGAGCGGTATCAATGATGCTTTCCGCTGGAATGACAAACAGCCTGAGATTCT	2615
Qy	725	uGlyIleGlnProThrLeuGlnLysProProAsnGlnProProValSerIleTrpLeuIleVal	745

Dd	2618	GGGGAATACAGCCAACTTGGACTCCTAACCAACCCTTGTTCCAATAGCGTAATTGG	267
Qy	745	lPheglYalValMetGlYalIleValIleValIgJlyIleValIIleLeuIIePheThrGlyIl	765
Dd	2678	TTTTGAACTGTGTGATGGAGTAGTAGTGTTGGCATGTGCATTCGTGATCTTCACTGGGAT	273
Qy	765	eArGaSPaRgLySLySLySaSnLysAlaAArgSerGIyGLuaSNProTYrNASeRIleAs	785
Dd	2738	CAGAGATCGAAGAAGAAAAATAAGACAAGAGGAGAAAATCCATTATGCCCTCATCGA	279
Qy	785	pLSerLySgLIyGLuaSNaSPnProGIYPheGlNaSThrASpAPyAlGIthrSerPh	805
Dd	2798	TATTAGCAAAGACAAATAATATCCAGCATTCCAAAAACACTGATGATGTTGACACTCTT	285
Qy	805	e 805	
Dd	2858	T 2858	
RESULT 6			
US-10-140-470-71			
; Sequence 71, Application US/I0140470			
; Publication No. US20030022331A1			
GENERAL INFORMATION:			
APPLICANT: Baker, Kevin P.			
APPLICANT: Beresini, Maureen			
APPLICANT: DeForge, Laura			
APPLICANT: Desnoyers, Luc			
APPLICANT: Filvaroff, Ellen			
APPLICANT: Gao, Wei-Qiang			
APPLICANT: Gerritsen, Mary E.			
APPLICANT: Goodard, Audrey			
APPLICANT: Godowski, Paul J.			
APPLICANT: Gunney, Austin L.			
APPLICANT: Sherwood, Steven			
APPLICANT: Smith, Victoria			
APPLICANT: Stewart, Timothy A.			
APPLICANT: Tumas, Daniel			
APPLICANT: Watanabe, Colin K			
APPLICANT: Wood, William			
APPLICANT: Zhang, Zemin			
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
FILE REFERENCE: P3330RIC160			
CURRENT APPLICATION NUMBER: US/10/140,470			
CURRENT FILING DATE: 2002-05-06			
Prior Application removed - See Palm or File Wrapper			
NUMBER OF SEQ ID NOS: 550			
SEQ ID NO 71			
LENGTH: 3732			
TYPE: DNA			
ORGANISM: Homo Sapien			
US-10-140-470-71			
Alignment Scores:			
Pred. No.: 0 Length: 3732			
Score: 4142.00 Matches: 802			
Percent Similarity: 85.33% Conservative: 1			
Best local Similarity: 85.23% Mismatches: 2			
Query Match: 96.53% Indels: 138			
Dd: 9 Gaps: 1			
US-09-978-385-2 (1-805) x US-10-140-470-71 (1-3732)			
Qy	1	MeterSerSerSerTriPLeuLeuSerLeuValAlaValThraAlaInserThr	20
Dd	40	ATGTCACACTTCTTCTGCGCTCTCTCACGCTTGTTCTTAACTGCTGCTCAGTCCACC	99
Qy	21	IleGluGlnAlaIaLYSLyThrPheLeuASPlySPheaSnHISgluaIaGluasPreuPhe	40
Dd	100	ATTGAGGAAAGCGCCAGACACTTTTGGACAAGTTTAAACCAAGCAAGCGCAAGACCTGTTC	159
Qy	41	TyrGlnSerSerLeuAlaSerTrTPAsnTYrAsnThraSnIleThrgIuGluasNValGln	60

```

Db 160 FTCAAGTCACTGCTTGGAAATATTAACCAATATTACTGAAAGAGATGTCCAA 219
QY 61 AsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 220 AACATGATTAATGCTGGGCAAAATGCTGCTTTTAAAGAACAGCTCCACTGCTCC 279
QY 81 GlnMetLysProLeuGlnGluLeuGlnAsnLeuThrValLysGlnLeuGlnAlaLeu 100
Db 280 CAAATGTATCCACTACAGAAATTCAGATTCACAGTCACAGCTTCACCTGAGGCTCTT 339
QY 101 GlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrIleLeu 120
Db 340 CAGCAAAATGGGCTTCAGTGTCTCAGAGAACAGCAACCAAGCGTTGAAACAATTTCTA 399
QY 121 AsnThrMetSerThrIleLysSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 400 AATACAAATGAGCACATCTACAGTACTGGAAAGTTTGTAACTCCAGATTAATCCACAAAGAA 459
QY 141 CysLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 460 TCCTTATTAATCTGAACCAAGTTTGAATGAATATGCAACAGTTTATGACTTACAAATGAG 519
QY 161 ArgLeuTPalaTPrpGluSerTPrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
Db 520 AGGCTCTGGGCTTGGGAAAGCTGAGATCTGAGTCCGCAACGAGCTGAGGCCATTATAT 579
QY 181 GlnGluTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspTyrGly 200
Db 580 GAAGAGTATGTGCTCTTAAATAATGATGTCACAAAGCAAAATCAATTATGAGAGCTATGGG 639
QY 201 AspTyrTPrpArgLysAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
Db 640 GATTATTTGGAGAGGAGACTATGAATTAATGGGAGTGTGCTATGACTTAACCCGCGCC 699
QY 221 GlnLeuIleGluAspValGlnHisThrPheGlnGluIleLysProLeuTyrGlnHisLeu 240
Db 700 CAGTGTATGGAAGATGTGGAACATACCTTTGAGAGATTAACCAATATATGAAACATCTT 759
QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 760 CATGCTTATGTAGAGGCAAAAGTTGATGATGATGCTTATCTCTATATCAAGTCCAAATTTGA 819
QY 261 CysLeuProAlaHisLeuLeuGlnLysPheTPrpGlyArgPheTPrpHisLeuTyrSer 280
Db 820 TGCCCTCCCTGCTCATTTCTGCTGATGATGTGGGTATGATTTTGACAAATCTTACTCT 879
QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 880 TTGACAGTTCCCTTTGGACAGAAACCAAAACATAGATTTACTGATGCAATGTGTGACACAG 939
QY 301 AlaTPrpAspAlaGlnArgIlePheLysGlnAlaGluLysPhePheValSerValGlyLeu 320
Db 940 GCCTGGGATGACAGACAGATATTTCAAGAGAGCCGAGAAAGTCTTGTATCTGTTGGTCTT 999
QY 321 ProAsnMetThrGlnGlyPheTPrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
Db 1000 CCTATATGACTCAAGAGATCTGGGAAATTCATGCTAAAGCAAGCCAGAAATGTTCTAG 1059
QY 341 LysAlaValLysHisProThrAlaTPrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360
Db 1060 AAAGAGCTCCCATCCACAGCTTGGAGACCTGGGAAAGCGAGCTTCAGAGATCTTATATG 1119
QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnIleGln 380
Db 1120 TGCACAAAGGTGACATGAGAGACTTCTGTACACACTCATCTAGATGAGGAGGCAATATCCAG 1179
QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400
Db 1180 TATGATATGCAATATGCTGACAAACCTTTCTCTAAGAAATGAGAGCTAATGAAAGATTC 1239
QY 401 HisGlnAlaValAlaGlyGlnIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420

```

---

```

Db 1240 CATCAACCTGTTGGGAAATCATGCTACTTCTGACCCACACCTTAACATTTAAATCC 1299
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGlnIleAsnPheLeu 440
Db 1300 ATTGCTCTTGTGACCCGATTTTTCAGAGAGACATGAAACAGAAATTAACCTTCGTCTC 1359
QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTPrpArg 460
Db 1360 AAACAGACACTCACAGATTTGGGACTCTGCTCATTTACTTACATGTTAGAGAGTGGAGG 1419
QY 461 TrpMetValPheLysGlyGluIleProLysAspGluTrpMetLysTPrpGluMet 480
Db 1420 TGGATGCTCTTAAAGGGGAAATTTCCAAAGACCAAGTGGATGAAAGTGGTGGAGATG 1479
QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500
Db 1480 AAGCAGAGATAGTGGGGGTGGAGAACCTGTGCCCATGATGAAACATCTGTGACCC 1539
QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu 520
Db 1540 GCATCTGTGCCATGTTTCTGATGATTAATCTCATTTCTGATTAATCAACAGGACCTT 1599
QY 521 TyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaAlaLysHisGlnGlyProLeuHis 540
Db 1600 TACCAATTCAGATTTCAAGAGACACTTTGTCAAGCAGCTTAACATGAAAGCCCTGTGAC 1659
QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe----- 555
Db 1660 AAATGTACATCTCAAACTCTACAGAAAGCTGACAGAAACTGTT -GTAAAGAAATACCTCA 1718
QY 555 ----- 555
Db 1719 AAATGTTGAACCTCTCTAGTATTAATCAATTTCCATGCTAGGTTGATTTGG 1778
QY 555 ----- 555
Db 1779 ATTTCTTGTGTTAAAGAAATTTTATGCGCTCAAAATGCTCATTTTACAAACANA 1838
QY 555 ----- 555
Db 1839 CATTAAATTTGTGTGACAGACAGAACCTAGACCATCAACAAATTTGGTGGCCACTCTT 1898
QY 555 ----- 555
Db 1899 TTCTCCCTATCAATACAGACCCCTCTCTCTGTAATTTGGAAGAAAGAGCGGTTTAG 1958
QY 555 ----- 555
Db 1959 GGTGAATATATCTGTTAATATGATCATTTCTTATCTGCGCAGAACAAATTTAGCCA 2018
QY 555 ----- 555
Db 2019 GTCAAGAGAGAAACCATAGATCATGATGTAATATATATGACATCTGAAACCCCTCA 2078
QY 556 ----- 556
Db 2079 AAAGCCCTGAGACCCCTTTTGTGTAGCATATCTCTAGAGCTTGGAAATTCAGAACCC 2138
QY 566 TrpThrLeuAlaLeuGlnAsnValValGlyAlaLysAsnMetAsnValArgProLeuLeu 585
Db 2139 TGGACCTTAGCATTTGAAATGTTGTAGAGAGC -AAGAACATGAAATGAGGCGACAGTTC 2197
QY 586 AsnTyrPheGlnProLeuPheThrTPrpLeuLysAspGlnAsnLysAsnSerPheValGly 605
Db 2198 AACTACTTTGAGCCCTTATTTACTGCTGGAAGAGCAACAAAGAAATCTTTTGGGGA 2257
QY 606 TrpSerThrAspTPrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeuLys 625
Db 2258 TGGAGTACGAGACTGAGAGTCCATATGACAGACCAAAACATCAAAAGTGAAGATTAAGCTTAA 2317
QY 625 sSerAlaLeuGlyAspLysAlaTyrGluTPrpAsnAspAsnGlnMetTyrLeuPheArgSe 645
Db 2318 ATCAGCTCTTGGAGATAAAGCATATGAAATGAAAGCAATGAAATGATACCTGCTCGATC 2377

```

QY	645	ISSVAlAAlATyAlAlMeArAgllrPhrHeuLysAlLysAsnGlnMeIleuPh	665
Db	2378	ATCTGTTGCATAGCCTATGAGCGACACTTTTAAAGATAAATAATCAGATATCTCTTT	2433
QY	665	eGlyGluGluAspAlaArgValAlaAsnLeuLysProAqIleSerPheAsnPhaPheVa	685
Db	2438	TGGGGAGAGAGAGTGGAGATGGCTAATTTGAAACCAAGATCTCCTTTAAATTTCTTGT	2492

QY	685	lThrala	prolys	snval	serasp	llelle	Proang	Thrgu	ValGul	lysala	lear	705
Db	2498	CACTGCA	CTAAA	ATGTGT	CTGATA	CACTTC	CTAGAC	TGAAGT	TGAAGG	CCATCAG	2555	

Qy 725 ugiyllelnProthleucllyproproksnclnProProvalSerlletrpleullela 745  
|||||  
|||  
Db 2618 GGGGATACACCAACTTGACCTCTTAACCAAGCCCCGTTCATATGGCGTATTGT 2677

Qy	765	eA	rG	sP	aR	gI	sP	rG	L	sY	sL	sY	sA	sL	aA	rG	S	e	r	G	L	u	a	n	P	r	o	t	r	a	L	a	s	e	r	I	e	a	s	785
Db	2738	C	A	G	A	G	A	T	C	G	A	A	G	A	A	A	T	A	A	G	C	A	G	A	G	T	G	G	A	A	A	A	T	C	C	T	T	A	T	2797

785 pleserlysglyuasnasnprcglypneinasthraspaspvalginttrserpn 805  
 Db 2798 TATTAGCAAGAGAGAAATATATCCAGGATTCCAAAACACTGATGATGTTCCAGACTCCTT 285

27	2858	—	2858
Db	2858	T	2858

US-10-175-746-71  
; Sequence 71, Application US/10175746  
; Publication No. US20030027270A1  
; GENERAL INFORMATION:

```

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P33301RC353
CURRENT APPLICATION NUMBER: US/10/175,746
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
SEQ ID NO 71
LENGTH: 3732

```

Pred. No.:	0	Length:	3732
Score:	4142.00	Matches:	802
Percent Similarity:	85.33%	Conservative:	1
Best Local Similarity:	85.23%	Mismatches:	2
Query Match:	96.53%	Indels:	138
DB:	9	Gaps:	1
US-09-978-385-2 (1-805)	x	US-10-175-746-71 (1-3732)	

QY 321 ProasmetInrGInglyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340  
 |||||  
 Db 1000 CCTATATAGACTCAAGAGATTCCTGGGAAAATTCATGCTAACGACGACCAGAAATGTTCCAG 1059  
 QY 341 LysAlaValCysHisProThrAlaTirPaspLeuGlyLysGlyAspPheArgLleLeuMet 360  
 |||||  
 Db 1060 AAGACAGCTCTCCATCCACACGCTTGGAGCTGGGAAAGGCGACTTCAGAGATCCTTAG 1119  
 QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisHisGluMetGlyHisLleGln 380  
 |||||  
 Db 1120 TGCACAAAGGTCACAAATGAGAGACTTCCTGACACGCTCATCATGATGAGGGCATATCCAG 1179  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyLysPhe 400  
 |||||  
 Db 1180 TATATATAGCATATGCGACCAACCTTTCTGCTAAGAAATGAGCATTAATGAGGATTC 1239  
 QY 401 HisGluAlaValGlyGluLleMetSerLeuSerAlaAlaThrProLysHisLleLysSer 420  
 |||||  
 Db 1240 CATGAGCTGTGGGAAATCATGCTACTTCTGACGCCACACCTAAGCATTTAAAAATCC 1299  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluLleAsnPheLeuLeu 440  
 |||||  
 Db 1300 ATGCTGCTCTGACCCGATTTTCAGAGACACAAATGAAACAGAAATTAACCTCTCTC 1359  
 QY 441 LysGlnAlaLeuThrLleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460  
 |||||  
 Db 1360 AAACAGCAGCTCAGCATTTGTTGGGACTCTGCCATTACTTACATGTTAGAGAACTGAGG 1419  
 QY 461 TrpMetValPheLysGlyGluLleProLysAspGlnThrPheLysLysTrpTrpGluMet 480  
 |||||  
 Db 1420 TGGATGGCTTTTAAAGGGAAATTCACAAAGACAGTGAGTGAAGAAAGTGGTGGAGATG 1479  
 QY 481 LysArgGluLleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500  
 |||||  
 Db 1480 AAGGAGAGATAGTTGGGGTGGTGAACCTGTGCCCATGATGAACATACCTGAGACCC 1539  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheLleArgTyrThrArgThrLeu 520  
 |||||  
 Db 1540 GCATCTCTGTTCCATGTTCTTGATGATGATTCATTCATTCGATATTCACAAAGACCTT 1599  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaLysHisGluGlyProLeuHis 540  
 |||||  
 Db 1600 TACCAATTCAGTTTCAAGAAAGCACTTGTCAACGACCTAAACCTGAAGGCCCTCTGCAC 1659  
 QY 541 LysCysAspLleSerAsnSerThrGluAlaGlyLysLysLeuPhe----- 555  
 |||||  
 Db 1660 AATGTGACATCTCAAACTCTACAGAAAGCTGGACAGAAACTGTT-GTAAAGAACTACTCA 1718  
 QY 555 ----- 555  
 Db 1719 AATGTGAACTCTCCTAGTATTCAGTATTACTATTCCATGCTAGGTTGTATTG 1778  
 QY 555 ----- 555  
 Db 1779 ATTTCTTTGTTTAAAGAAATTTTATGGCTCAAAATGTCTCATTTTACAACCCAA 1838  
 QY 555 ----- 555  
 Db 1839 CATTTAATTTGTGTCAGACAGAGACTAGACCATACAAATTTGGGTGGCCACCTCTT 1898  
 QY 555 ----- 555  
 Db 1899 TTCTCCTATCATTAACAGACCCCTCTCTCTGTGTAATTGGAAGAAAGAGGGTTAG 1958  
 QY 555 ----- 555  
 Db 1959 GGTGGAATATATCTGTTAATATGCAATCTTTCTTATCTGCAGAAACAAATTTAGCCAA 2018  
 QY 555 ----- 555  
 Db 2019 GTCAAAAGAAAGAACCATAGATCATAGTAAATATATGTACATCTGGAACCCCTCAA 2078  
 QY 556 -----AsnMetLeuArgLeuGlyLysSerGluPro 565

Db 2079 AAGGCCCTGAAACCCCTTTTGTGTAGCAATATGCTAGGCTTGAAATCAGAACCC 2138  
 |||||  
 QY 566 TrpThrLeuAlaLeuGluAsnValAlaGlyAlaLysAsnMetAsnValArgProLeuLeu 585  
 |||||  
 Db 2139 TGGACCTTACATTTGAAAATGTTGTAGGAGC-AAACATCATAAATGTAAGGCCACTGTCC 2197  
 QY 586 AsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGly 605  
 |||||  
 Db 2198 AACTACTTTGAGCCCTTATTACTTGCTGCTGMAAGACCAACAAAGATTTCTTTGTGGCA 2257  
 QY 606 TrpSerThrAspTrpSerProTyrAlaAsp-GlnSerLleLysValArgLleSerLeuLys 625  
 |||||  
 Db 2258 TGGAGTACCGAGCTGAGTGCATATGACAGACCCAAAGCATCAAAAGTGAAGATTAAGCTAAA 2317  
 QY 625 sSerAlaLeuGlyAspLysAlaTyrLleThrPaspAsnAspAsnGluMetTyrLeuPheArg 645  
 |||||  
 Db 2318 ATCAGCTCTTGGAGATTAAGCATATGAAATGAAACGCAATGAAATGTACCTCTCCGATC 2377  
 QY 645 rSerValAlaTyrAlaMetArgGlnThrPheLeuLysValLysAsnGluMetLleLeuPhe 665  
 |||||  
 Db 2378 ATCTGTGCATATGCTATGAGCAGTACTTTTAAAGTAAAAATCAGATATTTCTTTT 2437  
 QY 665 eGlyGluGluAspValArgValAlaAsnLeuLysProArgLleSerPheAsnPhePheVal 685  
 |||||  
 Db 2438 TGGGAGAGAGATGTGCGAGTGCCTAATTGAAACCAAGATCTCCTTAATTTCTTGT 2497  
 QY 685 LThrAlaProLysAsnValSerAspLleLeuProArgThrGluValGluLysAlaLleArg 705  
 |||||  
 Db 2498 CACTGCACCTTAAATAATGTCTGATATCATTCCTTGAACGTGAAGTGAAGGAGCATCAG 2557  
 QY 705 gMetSerArgSerArgLleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLe 725  
 |||||  
 Db 2558 GATGTCCCGAGCGCATATGATGATGCTTCCGCTGCAATGACAAACAGCTGAGATTTCT 2617  
 QY 725 uGlyLleGlnProThrLeuGlyProProAsnGlnProProValSerLleTrpLeuLleVal 745  
 |||||  
 Db 2618 GGGGATACAGCCAACTGTGACCTCTCAACAGCCCTGTTCCATGTGCTGATGTTGT 2677  
 QY 745 lPheGlyValAlaMetGlyValLleValAlaGlyLleValLleLeuLlePheThrGlyL 765  
 |||||  
 Db 2678 TTTTGGATGTGATGAGGAGTATGATGCTGCTGATGCTGATGCTGATGCTGATGCTG 2737  
 QY 765 eArgAspArgLysLysLysAsnLysAlaArgSerGlyGluAsnProTyrAlaSerLleAs 785  
 |||||  
 Db 2738 CAGAGTTCGGAAGAGAAATTAAGCAAGAGTGGAGAAATCTTATGCTCCATCGA 2797  
 QY 785 pLleSerLysGlyGluAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPhe 805  
 |||||  
 Db 2798 TATTACCAAGAGGAAATAATTCAGAGATTCCAAAACACTGATGATGTTCAACACTCCT 2857  
 QY 805 e 805  
 Db 2858 T 2858  
 |||||

## RESULT 8

US-10-176-918-71

Sequence 71, Application US/10176918

Publication No. US20030027275A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel  
 APPLICANT: Matanabe, Colin K  
 APPLICANT: Wood, William  
 APPLICANT: Zhang, Zemin  
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 TITLE OF INVENTION: ACIDS ENCODING THE SAME  
 FILE REFERENCE: P333081382  
 CURRENT APPLICATION NUMBER: US/10/176,918  
 CURRENT FILING DATE: 2002-06-20  
 Prior Application removed - See File Wrapper or Palm  
 NUMBER OF SEQ ID NOS: 550  
 SEQ ID NO 71  
 LENGTH: 3732  
 TYPE: DNA  
 ORGANISM: Homo Sapien  
 US-10-176-918-71

## Alignment Scores:

Pred. No.: 0 Length: 3732  
 Score: 4142.00 Matches: 802  
 Percent Similarity: 85.338 Conservative: 1  
 Best Local Similarity: 85.238 Mismatches: 2  
 Query Match: 96.538 Indels: 138  
 Gaps: 1

US-09-978-385-2 (1-805) x US-10-176-918-71 (1-3732)

QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThr 20  
 DB 40 ATGTCAAGCTCTTCCGCTCTCTCCACCTGTGCTGTAAGCTGCTCAGTCCACC 99  
 QY 21 IleglGluGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGluAspLeuPhe 40  
 DB 100 ATTTAGGAGACAGGCCAAGCATTTTGGACAAATTACACGAGAACCCGCAACCTGTTC 159  
 QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTrpAsnIleThrGluGluAsnValGln 60  
 DB 160 TATCAAAAGTTCACCTCTCTTCTGGAATTTAAACCAATATTACTGAAGAGATGTCCAA 219  
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla 80  
 DB 220 AACATAAATATGCTGGGACAAATGCTGCTTTTAAAGAACAGTCCACACTGGCC 279  
 QY 81 GlnMetTrpProLeuGlnLysIleGlnAsnLeuThrValLysLeuGlnLeuAlaLeu 100  
 DB 280 CAAATGTATCCACTACAAATAATTCGATCTCACAGTCAAGCTTCAGCTGAGGCTCTT 339  
 QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120  
 DB 340 CAGCAAAATGGCTCTCAGTCTCTCAGAACAGAACGAGGTTGAAACACAATTTCTA 399  
 QY 121 AsnThrMetSerThrIleLeuSerThrLysValCysAsnProAspAsnProGlnGlu 140  
 DB 400 AATACATATGACACCATCTACAGTACTGGAAAAGTTGTAAACCCAGATATATCCACAAGA 459  
 QY 141 CysIleuLeuLeuGluProGluLeuAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160  
 DB 460 TGGTATTACTGTAACCAAGTTGATGAATATATGCAACAGTTTAACTACATGAG 519  
 QY 161 ArgLeuTrpAlaTrpLeuSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTrp 180  
 DB 520 AGGCTCTGGGCTGGAAAAGCTGGAGATCTGAGTGGGCAACAGCTGAGGCAATTATAT 579  
 QY 181 GlnGluTrpValIleuLeuAsnGlnMetAlaArgAlaAsnHisTrpGluAspTyrGly 200  
 DB 580 GAAGAGATATGTGCTTGAATAATGATGGCAAGAGCAATATATATGAGGATATGGG 639  
 QY 201 AspTyrTrpArgLysPyrGluValAsnGlyValAspGlyTyrAspTyrSerArgLys 220  
 DB 640 GATTATTTGGAGAGACATATGATGAATGGGTAGATGGCTATGACACGCGCGGC 699  
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTrpGluHisLeu 240

DB 700 CAGTTGATGAGATGTGGAACATACCTTTGAAGAGATTAACCATATATGAAACATCTT 759  
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerLysIleSerProIleGly 260  
 DB 760 CATCCCTATGTGAGGGGCAAGTGTATGAATGCCATATCTCCATATACATCCCAATTTGA 819  
 QY 261 CysLeuProAlaHisLeuLeuGluLysPheMetTrpGluValArgPheTrpPheAsnLeuTrpSer 280  
 DB 820 TGGCTCCCTGCTCATTTGCTTGGTATGATGTGGGTGATTTTGGACAAATCTGTACTCT 879  
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 DB 880 TTGACAGTCCCTTGGACAGAAACCAACATAGATGTACTGATGCAATGAGTGAGACACAG 939  
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320  
 DB 940 GCCGTGGATGCAAGAAATATTAAGAGAGCCAGAAAGTTCTTGTATCTGTGGTCTT 999  
 QY 321 ProAsnMetTrpGlnGlyPheTrpGluAsnSerMetLeuThrAspProGluAsnValGln 340  
 DB 1000 CCTATATGACTCAAGGATTTCTGGGAAAATTTCCATGCTAACGAGCCAGGAAATGTTTCAG 1059  
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 DB 1060 AAACGAGCTGTCATCCACAGCTTGGGAGCTGGGAGGAGGAGCATTCAGATCTTATG 1119  
 QY 361 CysThrLysValIleThrMetAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380  
 DB 1120 TGCACAAAGGAGACAAATGAGAGACTCTTGACACTATATGATGATGAGGGGCAATTCACAG 1179  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400  
 DB 1180 TATGATATGCGATATGCTGGCAACACTTTCTGCTAAGAAATGAGGATATGAAAGATTC 1239  
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
 DB 1240 CATGAAGCTGTGGGGAATCATCTCACTTCTCCACCCACCTTAAGCATTAATTAATCC 1299  
 QY 421 IleglyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440  
 DB 1300 ATTTGCTCTCTGTCACCCGATTTTCAGAGACATATGAACAGAAATAACTCTGCTC 1359  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460  
 DB 1360 AAACAGACACTCAGATTTGTTGGAGCTGCGACTGCAATTTACTTAACATGTAGAGAGTGAGG 1419  
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGluIleTrpMetLysTrpTrpGluMet 480  
 DB 1420 TGGATGCTCTTAAAGGGGAAATTTCCAAAGACAGTGGATGAATAAGTGTGGGAGATG 1479  
 QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500  
 DB 1480 AAGCAGAGATATGTTGGGCTGTGGAACTGTGGCCCATGATGAACATATCTGACCCC 1539  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520  
 DB 1540 GCATCTGCTGCCATGTTTCTGATGATATACCATTTCTGATATATGACAAAGAGACCTT 1599  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaIleLysHisGluGluLysProLeuHis 540  
 DB 1600 TACCAATTTCCAGTTTCAAGAGCACTTGTTCACACCACTAAACATGAAGGCCCTGTGCAC 1659  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPhe 555  
 DB 1660 AAATGTGACATCTCAAACTCTACAGAAAGCTGGACAGAAACTGTT-GTAAACAATACCTCA 1718  
 QY 555 555  
 DB 1719 AAATGTGAACCTCTCTAGTATTCAGTATATCTCATTTCCAGGCTAGTTGTATTTG 1778  
 QY 555 555  
 DB 1779 ATTTCTTTGTTCTAAAGAAATTTATGAGCTCAAAATGCTCCTCATTTTAAACCAAA 1838





QY 141 CysLeuLeuLeuGluProGlyLeuAsnGluMetAlaAsnSerLeuAspTyrAsnGlu 160  
 |||||  
 Db 460 TGCCTATTAATGAAACAGGTTGATGAATAATGCAACAGTTTGACATCAATGAG 519  
 QY 161 ArgLeuTrpAlaTrpGlySerTrpArgSerGlyValGlyGlnLeuArgProLeuTyr 180  
 |||||  
 Db 520 AGGCTCTGGGCTGGGAAAGCTGGAGATCTGAGTCGGCAAGCAGCTGAGGCATATAT 579  
 QY 181 GluGluTyrValValLeuLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200  
 |||||  
 Db 580 GAAGGATAGTGTCTTGAATAATGAGATGGCAAGCAAAATCATTAATGAGGACTATGG 639  
 QY 201 AspTyrTrpArgGlyAspTyrGlyValAsnGlyValAspGlyTyrAspTyrSerArgGly 220  
 |||||  
 Db 640 GATTATGAGAGAGAGACATATGAGTAATGGGTAGATGGCTATGACTACAGCCGGC 699  
 QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240  
 |||||  
 Db 700 CAGTTGATTGAGAGATGAGACATACCTTGAAGAGATTAACCATTAATGACATCTT 759  
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260  
 |||||  
 Db 760 CATGCTATGATGAGGCAAAAGTTGATGATGCTTCTCTATATCACTCAATGTGA 819  
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280  
 |||||  
 Db 820 TGCCTCCCTGCTCATTTGCTTGGTATGATGCGGTAGATTGTCGACAAATGTACTCT 879  
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 |||||  
 Db 880 TTGACAGTTCCCTTGGACAGAAACCAACATGATGTACTGATGCAATGGTGGACAG 939  
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu 320  
 |||||  
 Db 940 GCGTGGAGTGCACAGAAATATTCAGAGGCCGGAAGATCTTTGTATCTTTGCTCTT 999  
 QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340  
 |||||  
 Db 1000 CCTAATATGATGCTCAAGGATTTGGGAAATTCATGCTCAAGGAGCCAGAAATGTGAG 1059  
 QY 341 LysAlaValAlaCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 |||||  
 Db 1060 AAAGCAGTCTGCCATCCACAGCTTGGGACCTGGGGAAGGGGAGCTTCGAGTCCCTTAG 1119  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380  
 |||||  
 Db 1120 TGCACAAAGGTGACAAATGACGACCTTCTGACAGCTCATGATGAGATGGGCAATCCAG 1179  
 QY 381 TyrAspMetAlaTyrAlaIleGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe 400  
 |||||  
 Db 1180 TATGATATGGCATATGCTGCACAACTTTTCTGCTAAGAAATGAGACTAATGAGAGATTC 1239  
 QY 401 HisGluAlaValGlyGluIleMetSerLeuSerAlaIleThrProLysHisLeuLysSer 420  
 |||||  
 Db 1240 CATGAAGCTGTGGGAAATCATGCTACTTTCGACGCCACACTAAGCATTTAAATATCC 1299  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440  
 |||||  
 Db 1300 ATTGGTCTTCTGTCACCCGATTTTCAAGAGACATGAAACGAATAAATCTTCGCTGC 1359  
 QY 441 LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460  
 |||||  
 Db 1360 AAACAAGCACTCAGATGTGTGGACTCTGCCATTTACTTACTATGTAAGAAGTGGAG 1419  
 QY 461 TrpMetValPheLysGlyGluIleProLysAspGlnTrpMetLysLysTrpGluMet 480  
 |||||  
 Db 1420 TGGATGGTCTTTAAAGGGGAAATTCCTCAAGACAGCTGATGAAAAAGTGGGAGATG 1479  
 QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500  
 |||||  
 Db 1480 AACCGAGAGATAGTTGGGGGTGGGAACCTGTGCCCATGATGAACATACCTGTGACCCC 1539  
 QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520

Db 1540 GCATCTGTTCCATGTTTCTGATGANTACTCTATTCATTCGATATACAAAGAGACCTT 1599  
 QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis 540  
 |||||  
 Db 1600 TACCAATTCAGATTTCAGAAAGCACTTGTCAAGCAGCTTAACATGAAAGCCCTCTGAC 1659  
 QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGluIleLysLeuPhe 555  
 |||||  
 Db 1660 AAATGTACATCTCAAACTCTACAGAAAGCTGACAGAAAGTGT-CTAAGAAATACCTCA 1718  
 QY 555 555  
 Db 1719 AAATGTGAACCTCTCTAGTATTCAGATTAATCTCAATTCACAGCTAGGTTGTATTGG 1778  
 QY 555 555  
 Db 1779 ATTCTTGTGTTCAAAAAAATTTTATGGCCTCAAAATGCTCTCATTTACAACCCAA 1838  
 QY 555 555  
 Db 1839 CATTTAATTTGTGTCAGACAGAACCTAGACCATACAACAATTTGGGTGGCCACCTCTT 1898  
 QY 555 555  
 Db 1899 TTCTCCCTATCATTAACACAGCCCTCTCTCTGTGTAATTGGAAGAAAGAGCGTTAG 1958  
 QY 555 555  
 Db 1959 GGTGGAATATATCTGTTAATATGCAATCTTTCTTATCTCCAGAAACCAATTTAGCCA 2018  
 QY 555 555  
 Db 2019 GTCAAAAGAGAAACCATAGATCATATATATATGATCATCTGAAACCCCTCAA 2078  
 QY 556 556  
 Db 2079 AAGGCCCTGAACCCCTTTTGTGTAGCAATATGCTGAGGCTGGAAATATAGAAACCC 2138  
 QY 566 566  
 Db 2139 TGGACCTCAGATTTGGAAATAGTTGTAGAGAC-AAGAACATGATATGAGCCACCTGCTC 2197  
 QY 586 AsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSerPheValGly 605  
 |||||  
 Db 2198 AACTACTTTGAGCCCTTATTTACTCGCTGAAGACCAAGCAAGATCTTTTGTGGGA 2257  
 QY 606 TrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeuLys 625  
 |||||  
 Db 2258 TGGAGTACCGACTGAGTCCATATGACAGACCCCAAGCATCAAAAGTGAAGATTAAGCTTAA 2317  
 QY 625 sSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMetTyrLeuPheArgSe 645  
 |||||  
 Db 2318 ATCAGCTTTGGAGATTAAGCATATGATGAAGACAGACAAATGATGACTGTTCCGATC 2377  
 QY 645 rSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIleLeuPhe 665  
 |||||  
 Db 2378 ATCTGTTGCATATGCTATGAGGAGTACTTTTAAAGTAAATAATCAGATGATCTTTT 2437  
 QY 665 eGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheVal 685  
 |||||  
 Db 2438 TGGGAGGAGAGATGTCAGTGGCTAATTTGAAACCAAGATTCCTTAATTTCTTTGT 2497  
 QY 685 lThrAlaProLysAsnValSerAspIleIleProArgThrGluValGluLysAlaIleArg 705  
 |||||  
 Db 2498 CACTGACCTTAATAATGCTGTGATATCATTCCTAGAACTGAATGAAGGACCATCAG 2557  
 QY 705 gMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPheLe 725  
 |||||  
 Db 2558 GATGTCGGGAGCCGTATCAATGATGATGCTTCGCTGTGATGACAAACGCTAGAGTTCT 2617  
 QY 725 uGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTrpLeuIleVal 745

Db 2618 GGGGATACAGCCACTTGACCTGCTTAACACGCCCGCTGTTCCATGATGCTGATGT 2677  
 QY 745 IPhEGYVALIValMetGlyValIleValValGlyIleValIleIleIlePheThrGlyI 765  
 Db 2678 TTTTGGAGTTGATGGAGTGTAGTGGTGGCATTCGATCCGATCTTCATCGGGAT 2737  
 QY 765 eATGAPATGlyLysLysAsnLysAlaIleSerGlyLysAsnProTyrAlaSerIleAs 785  
 Db 2738 CAGAGATCCGAGAGAAATAAATTAACCAAGAGATGAGAGAAATCTTATGCTCCATGCA 2797  
 QY 785 pIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPh 805  
 Db 2798 TATTAACAAGAGAAATAAATTAACCAAGATTCACAAACACTGATGATGATGACACCTT 2857  
 QY 805 e 805  
 Db 2858 T 2858

## RESULT 10

US-10-137-865-71  
 ; Sequence 71, Application US/10137865  
 ; Publication No. US20030032155A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: Deforge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerlitsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Tumas, Daniel  
 ; APPLICANT: Watanabe, Colin K  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P330R1C154  
 ; CURRENT APPLICATION NUMBER: US/10/137,865  
 ; CURRENT FILING DATE: 2002-05-03  
 ; Prior Application removed - See Palm or File Wrapper  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 71  
 ; LENGTH: 3732  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-137-865-71

## Alignment Scores:

Pred. No.: 0 Length: 3732  
 Score: 4142.00 Matches: 802  
 Percent Similarity: 85.33% Conservaive: 1  
 Best Local Similarity: 85.23% Mismatches: 2  
 Query Match: 96.53% Indels: 138  
 DB: 9 Gaps: 1

US-09-978-385-2 (1-805) x US-10-137-865-71 (1-3732)

QY 1 MetSerSerSerTrpLeuLeuSerLeuValAlaValAlaThrAlaAlaGlnSerThr 20  
 Db 40 ATGTCAAGCTCTCTCTGCTCTCTCTCAAGCTTGTCTTAAGCTGCTCAAGTCCACC 99  
 QY 21 IleGlnGlnGlnAlaLysThrPheLeuAspLysPheAsnHisGlnAlaGlnAspLeuPhe 40  
 Db 100 ATTAGAGAACAGCCCAAGACATTTTGGACAGATTAAACACGAGCCGGAAGACCTTTC 159  
 QY 41 TyGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGlnGlnAsnValGln 60

Db 160 TATCAAAATTCCTGCTTCTTGGAAATTAATACACCAATATTACTGAAGAAATGCCAA 219  
 QY 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGlnGlnSerThrLeuAla 80  
 Db 220 AACATGAATATATGCTGGGGACAAATGCTCTCTTTTAAAGAAACGTCACACTTCC 279  
 QY 81 GlnMetLysProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnLeuAlaLeu 100  
 Db 280 CAAATGATTCACATCAAGAAATTCAGAAATCTCAGACGAAAGCTTCAGCTGACGCTCT 339  
 QY 101 GlnGlnAsnGlySerSerValLeuSerGlnAspLysSerLysArgLeuAsnThrIleLeu 120  
 Db 340 CAGCAAAATGGGCTTTCAGTCTCTCTCAGAAAGACAGCAAAACGTTGAACACAAATCTA 399  
 QY 121 AsnThrMetSerThrIleTyrSerThrGlyLysValLysAsnProAspAsnProGlnGln 140  
 Db 400 AATACAAATGAGACCATCTACAGTACGTAAGAAAGTTTGTAAACCCAGATATCCCAAGAA 459  
 QY 141 CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGln 160  
 Db 460 TGCTTATTACTTGAACCGAGGTTTGAATGAATTAATGGCAACAGTTTAAGACTACATGAG 519  
 QY 161 ArgLeuTrpPalaTrpGlnSerTrpArgSerGlnValGlyLysGlnLeuArgProLeuTyr 180  
 Db 520 AGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGCTGGCAAGCTGAGGCTATATAT 579  
 QY 181 GlnGlnTyrValValLeuLysAsnGlnMetAlaArgAlaAsnHisTyrGlnAspTyrGly 200  
 Db 580 GAAGAGTATGAGTCTCTTAAATAATGATGCGACAGCAAAATCTTATGAGAGATATGGG 639  
 QY 201 AspTyrTrpArgGlyAspTyrGlyLysValAsnGlyValAspGlyTyrAspTyrSerArgGly 220  
 Db 640 GATTAATGAGAGGAGACTATGAATTAATGGGAGATAGTCTATGACTACAGCCGCGGC 699  
 QY 221 GlnLeuIleGlnAspValGlnHisThrPheGlnGlnIleLysProLeuTyrGlnHisLeu 240  
 Db 700 CAGTTGATTAAGATGTGAACATCTCTTGAAGAGATTAACCATTAATGAACATCTT 759  
 QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260  
 Db 760 CATGCCATATGTAGGGCAAGTTGATGATGCCATCTCTCTATATCAGTCCCAATGGA 819  
 QY 261 CysLeuProAlaHisLeuLeuGlnLysAspMetTrpGlyArgPheTrpThrAsnLeuTyrSer 280  
 Db 820 TCCGACCCCGCTCAATTCGCTGTGTATGAGGGGTAGATTTTGACAAATCTGACTCT 879  
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 Db 880 TTGACAGTTCCTTTGGACAGAAACCAACATGATGTTACTGATGATGATGATGATGATGATG 939  
 QY 301 AlaTrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerValGlyLeu 320  
 Db 940 GCTTGAGATGCACAGATATATTCAGAGAGCGCGAAGTTCTTGTATCTGTTGGCTCTT 999  
 QY 321 ProAsnMetThrGlnGlyPheTrpGlnLysSerMetLeuThrAspProGlyAsnValGln 340  
 Db 1000 CCTAATATGACCAAGAGATTCGCGAAATTCATGCTACAGCAGCAGCAGGAAATGTTAG 1059  
 QY 341 LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 Db 1060 AAAGCAGTCTGCATCCACACACTTGAGCTGGGAGAGGCGGACTTACAGATCCTTATG 1119  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHisIleGln 380  
 Db 1120 TGCACAAAGATGACATGAGACATATTCGACAGCTCATCATGATGAGATGGGATATCCAG 1179  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPhe 400  
 Db 1180 TATGATATGCAATATGCTGACCAACCTTTTCTGCTAAGAAAGAGCAATTAAGATATC 1239  
 QY 401 HisGlnAlaValGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeuTyrSer 420  
 Db 1240 CATGAAGCTGTTGGGAAATCATCTCTTCTGAGCCACACTTAAGCATTTAAATCC 1299





```

Db 2079 AAGCCCTGAGACCCCTTTTGTGTAGCAATATGCTGAGGCTTGGAATAATCAGAACCC 2138
QY 566 TTTTThleuAlaIeugluAsnValValGlyAlaIeAsnMetAsnValArgProleu 585
Db 2139 TGGACCTTGACATGAGAAATGTTGAGAGC-AAGAACATGAAATTAAGCCACGCTC 2197
QY 586 AsnTyrPhegluProleuPheThrTrpLeuLysAspGluAsnLysAsnSerPheValGly 605
Db 2198 AACTACTTTGAGCCCTTATTTACCTGGCGGAAGACCAACAGAAATCTTTGGGGA 2257
QY 606 TrpSerThrAspTrpSerProTyrAlaAsp-GlnSerIleLysValArgIleSerLeu 625
Db 2258 TGGAGTACCGACTGAGATCATATGACAGCCAAACGATCAAAAGTGAATTAAGCCTAA 2317
QY 625 sSerAlaIeugluAspLysAlaIeTyrGluTrpAsnAspGluMetTyrLeuPheArg 645
Db 2318 ATCAGCTCTTGAGATTAAGCATATGAAAGAACGACAAATGAAATGCTTCCGATC 2377
QY 645 rSerValAlaIeTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMetIleu 665
Db 2378 ATCTGTGCAATATGATATGAGCAGACTTTTAAAGTAAAGTAAATCAGATGATCTTT 2437
QY 665 eGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePhe 685
Db 2438 TGGGAGAGAGATGTCGAGTGCCTAATTGAAACCAAGAAATCTCTTAATTTCTTGT 2497
QY 685 lThrAlaProLysAsnValSerAspIleIleProArgTrpGluValGluLysAlaIeAr 705
Db 2498 CACTGACCTGAAATATGTCTGATATCATCTCTGAACTGAAATGAAAGGCAATCG 2557
QY 705 gMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGluPhe 725
Db 2558 GAGTCCCGGAGCGCTATCATGATGATGCTTCCGTGATGACAAACAGCTAGATTTCT 2617
QY 725 uGlyIleGlnProThrLeuGlyProProAsnGlnProValSerIleTrpLeuIleVa 745
Db 2618 GGGGATACAGCCACACTTGGACCTCTTACACAGCCCTGTTCATATGCTATGT 2677
QY 745 lPheGlyValIleValMetGlyValIleValGlyIleValIleLeuIlePheThrGly 765
Db 2678 TTTTGGATGTTGATGGAGTGTATGATGTTGGCTTGCATTCATCTGATTTACGTGGAT 2737
QY 765 eArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnProTyrAlaSerIleAs 785
Db 2738 CAGAGATCGGAAGAGAAATAAACAAGCAAGAGTGAAGAAATTCCTATAGCTCCATGA 2797
QY 785 PileSerLysGlyLysAsnAsnProGlyPheGlnAsnTrpAspAspValGlnThrSerP 805
Db 2798 TATTACCAAGAGAAATAATCATCCAGATTCACAAACACTGATGATGTTCAAGACCTCT 2857
QY 805 e 805
Db 2858 T 2858

```

```

; APPLICANT: Matanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NOCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; PRIORITY DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 71
; LENGTH: 3732
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-71

Alignment Scores:
Pred. No.: 0 Length: 3732
Score: 4142.00 Matches: 802
Percent Similarity: 85.33% Conservative: 1
Best Local Similarity: 85.23% Mismatches: 2
Query Match: 96.53% Indels: 138
DB: 9 Gaps: 1

US-09-978-385-2 (1-805) x US-10-142-431-71 (1-3732)

QY 1 MetSerSerSerSerTrpLeuLeuSerLeuValAlaValThrAlaIleGlnSerThr 20
Db 40 ATGCAACCTCTTCTGCTCTCTCTCTCAGCCCTTGTCTGTAACCTGCTGCTCAGTCCACC 99
QY 21 lIleGluGluGlnAlaLysThrPheLeuAspLysPheAsnHisGluAlaGluAspLeuPhe 40
Db 100 ATTGAGAGACAGGCGCAAGACATTTTGGACAGATTTAACCAAGAGCGGAGACCTGTC 159
QY 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluGluAsnValGln 60
Db 160 TATCAAGTTCACCTGCTCTTCTTGAATTAATTAACCAATATTTCTGAAAGAAATGTCCAA 219
QY 61 AsnMetAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 220 AACATGAATTAATGCTGGGCAAAATGCTGCTCTTTTAAAGGAACGTCACACCTTGCC 279
QY 81 GlnMetTyrProLeuGlnIleGlnAsnLeuThrValLysLeuGlnIleGlnAlaLeu 100
Db 280 CAATATGATCCTACCAAGAAATTCAGATCTCAGATCTCAGCTCAGCTGACGCTCTT 339
QY 101 GlnGlnAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu 120
Db 340 CAGCAAAATGGGCTCTTCAAGTCTCTCAGAAAGACAAAGAGCAAGGTTGAACACAAATTC 399
QY 121 AsnThrMetSerThrIleTyrSerThrGlyValLysAsnProAspAsnProGlnGlu 140
Db 400 AATCAATAGACACACCTCTACGATCTGAGAAATGTTGTAAACCAATATTCACAAAGAA 459
QY 141 CysLeuLeuLeuGluProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 460 TGCCTATTACTTGAACACAGGTTTGAATGAATATATGCAAAACAGTTTAGCTCAATGAG 519
QY 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyr 180
Db 520 AGGCTCTGGGCTTGGAAGCTGAGATCTGAGCTCGGCAAGCAGCGTGAAGCCATATAT 579
QY 181 GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly 200
Db 580 GAAAGATATGTTGCTTGAAGAAATGAGATGGCAAGACCAATATCTATGAGACTATGGG 639
QY 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrAspTyrSerArgGly 220
Db 640 GATTATTTGAGAGAGACATGAAGTAAATGGGCTTAATGCTATGATGCTACACCCGCGC 659
QY 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240
Db 700 CAGTTGATTAAGATGAGTGAACATACCTTTGAAGAGATTAACCATATATGAAACATCTT 759

```

```

RESULT 12
US-10-142-431-71
; Sequence 71, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel

```

QY 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260  
 Db 760 CATGCTATGTGAGGGCAAAAGTTGATGAAATGCTATCTCTCTATATCCATTCGAATTTGA 819  
 QY 261 CysLeuProAlaHisLeuLeuGlyAspMetTrrpGlyArgPheTrrpHisLeuTyrSer 280  
 Db 820 TGGCTCCCTGGCTCATTTGGCTGTGATGTGGGTAGATTTTGGCAAAATCTGTACTCT 879  
 QY 281 LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln 300  
 Db 880 TTACAGTTTCCCTTTGGACAGAAACCAACATAGATGATTCATGCAATGGTGGACCCAG 939  
 QY 301 AlaTrrpAspAlaGlnArgIlePheLysGlnAlaGlnLysPheAspValSerValGlyLeu 320  
 Db 940 GCCTGGATGACACAGAAATATTCAGAGAGCGGAGAAATCTCTGTACTGTGTGTCTT 999  
 QY 321 ProAspMetThrGlnGlyPheTrrpGlyAsnSerMetLeuThrAspProGlyAsnValGln 340  
 Db 1000 CCTAATATGACTCAAGGATTTGGGAAATTCATGCTTACGAGACCCAGGAATGTTCAG 1059  
 QY 341 LysAlaValAlaCysHisProThrAlaTrrpAspLeuGlyLysGlyAspPheArgIleLeuMet 360  
 Db 1060 AAAGCAGTCTGCCATCCACAGCTTGGACCTGGGAGAGGCGACTTCAGGATCCCTTAG 1119  
 QY 361 CysThrLysValThrMetAspAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380  
 Db 1120 TGCACAAAGTGACAAATGACAGCACTTCGACAGCTCATGAGATGGGCGATATCCAG 1179  
 QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPhe 400  
 Db 1180 TATGATATGGCATATGCTCCACACACTTTTCTGCTAAGAAATGGAGCTAATGAAGATTC 1239  
 QY 401 HisGlnAlaValAlaGlyLysIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420  
 Db 1240 CATGAAGCTGTGGGAAATCATGTCACTTCTGACGCGACACCTGAGGATTTAAATTC 1299  
 QY 421 IleGlyLeuLeuSerProAspPheGlnGlnAspAsnGlnThrGlnIleAsnPheLeuLeu 440  
 Db 1300 ATTGGCTTCTGTCCACCCGATTTTCAGAAACATGAACAGAAATAACTCTCTGCTC 1359  
 QY 441 LysGlnAlaLeuThrIleValGlyLysIleThrLeuProPheThrTyrMetLeuGlnLysTrrpArg 460  
 Db 1360 AAACAGACCTCAGCATTTGTTGGACTGTGCACTTACTTACATCTTGAAGAGTGGAGG 1419  
 QY 461 TrpMetValPheLysGlyGlnIleProLysAspGlnTrrpMetLysLysTrrpGlnMet 480  
 Db 1420 TGGATGGCTTTAAAGGGGAATTTCCCAAGACCAAGTGGATGAAAAGTGGGAGATG 1479  
 QY 481 LysArgGlnIleValGlyValAlaGlnProValProHisAspGlnThrTyrCysAspPro 500  
 Db 1480 AAGCGAGAGATAGTTGGGGTGGTGAACCTGTGCCCATGATGAACATACGTGATCCCC 1539  
 QY 501 AlaSerLeuPheHisValSerAsnAspLysPheIleArgTyrTrrpThrArgThrLeu 520  
 Db 1540 GCATCTCTGTTCCATGTTTGTGATTAATCACTTCACTTCGATTTACACAAAGGACCTT 1599  
 QY 521 TyrGlnPheGlnPheGlnGlnAlaLeuCysGlnAlaIleAlaLysHisGlnGlyProLeuHis 540  
 Db 1600 TACCAATTCATTCATTCAGAAAGCACTTTGTCTCAAGCAGCTAAACATGAAGGCCCTGTGAC 1659  
 QY 541 LysCysAspIleSerAsnSerThrGlnAlaGlnLysLeuPhe----- 555  
 Db 1660 AAATGTGACATCTCAAACTCTACAGAGCTGGACAGAAAGCTT- GTAAGAAATATCTCA 1718  
 QY 555 ----- 555  
 Db 1719 AAATGTGAACCTCTCCTAGATTCAGATTAATCACTTTCATGCTAGCTTGTATTTG 1778  
 QY 555 ----- 555  
 Db 1779 ATTCTTTGTCTAAAGAAATTTTATGGCTCAAAATGTCTCATTTTACAAACCAA 1838

QY 555 ----- 555  
 Db 1839 CATTTAATTTGTGTGACAGACAGAACTAGACCATACAAATTTGGTGGCCACCTTT 1898  
 QY 555 ----- 555  
 Db 1899 TTTCCTCATATACATACAGCCCTCTCTCTCTGTATTTGGAAGAAAGCGGTTTAC 1958  
 QY 555 ----- 555  
 Db 1959 GGTGAATATATCTGTATATGATCTTTCTTATCTGCCAGAACCAATTTAGCCAA 2018  
 QY 555 ----- 555  
 Db 2019 GTCAAGAGAAAGAAACCATAGATCATAGATGTAAATATATGTACATCTGGAAACCTCAA 2078  
 QY 556 -----AsnMetLeuArgLeuGlyLysSerGlnPro 565  
 Db 2079 AAGGCCCTGAACCCCTTTTGTGTGATGCAATATGCTGAGAGCTTGGAAATACGAACCC 2138  
 QY 566 TrrpThrLeuAlaLeuGlnAsnValAlaGlyAlaLysAsnMetAsnValArgProLeuLeu 585  
 Db 2139 TGGACCTTAGCATTTGGAATAATGTTGTAGGAGC- AAGAAACATGAATGTAAAGCCACTGCTC 2197  
 QY 586 AsnTyrPheGlnProLeuPheThrTrrpLeuLysAspGlnAsnLysAsnSerPheValGly 605  
 Db 2198 AACTACTTTGAGCCCTTATTTTACTGCTGCAAGACAGACAGAAAGAAATCTTTGTGGGA 2257  
 QY 606 TrrpSerThrAspTrrpSerProTyrAlaAsp- GlnSerIleLysValArgIleSerLeuLys 625  
 Db 2258 TGGAGTACCGAGCTGAGTCCATATGACAGACCCAAAGCATCAAAATGAGATTAAGCCCTAA 2317  
 QY 625 sSerAlaLeuGlnLysPlyAlaTyrGlnTrrpAspAsnGlnMetTyrLeuPheArgse 645  
 Db 2318 ATCAGCTCTTGTGAGATTAAGCATATGAATGGAACAGACAAATGAATGTACCTGTTCCGATC 2377  
 QY 645 rSerValAlaTyrAlaMetArgGlnTrrpPheLeuLysValLysAsnGlnMetIleLeuPhe 665  
 Db 2378 ATCTGTTGCATATGCTATGAGGCGAGTACTTTTAAAGTAAATATAGATGATCTTTT 2437  
 QY 665 eGlyGlnGlnAspValArgValAlaAsnLeuLysProArgIleSerPheAsnPhePheVal 685  
 Db 2438 TGGGAGAGGAGATGTGGAGTGGCTAATTTGAACCAAGATCTCTTTATTTCTTTGT 2497  
 QY 685 lThrAlaProLysAsnValSerAspIleIleProArgThrGlnValGlnLysAlaIleArg 705  
 Db 2498 CACTGCACCTTAAATAATGTGTGATATCATTTCTTGAAGACGAAAGTTCAAAAGGCCATCAG 2557  
 QY 705 gMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLeuGlnPheLe 725  
 Db 2558 GATGTCCCGGAGCGGTATCAATGATGTCTTCCGTCTGAATGACACAGCTTAGAGTTTCT 2617  
 QY 725 uGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTrrpLeuIleVal 745  
 Db 2618 GGGGATACAGGCAACACTGTGACCTCTTAACCAAGCCCCGTTTCCATATGCGTGAFTGT 2677  
 QY 745 lPheGlyValAlaMetGlyValIleValAlaGlyIleValIleLeuIlePheThrGlyI 765  
 Db 2678 TTTTGGAGTTGTGATGGAGGTGATGATGCTTGGCATTTGCTATCTGATCTTCACTGGGAT 2737  
 QY 765 eArgAspArgLysLysLysAsnLysAlaArgSerGlnGlnAspProTyrAlaSerIleAs 785  
 Db 2738 CAGAGATCGSAGAGAAAGAAATTAAGCAGAAAGTGGAGAAATCTTATGCTCCATCA 2797  
 QY 785 pIleSerLysGlyGlnAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerPhe 805  
 Db 2798 TATTAGCAAGAGAAATAATATCCAGAGATTCAAAACATGATGATGTTCCAGACCTCCTT 2857  
 QY 805 e 805  
 Db 2858 T 2858  
 RESULT 13



Dd	460	TCCTTATTACTTGAACCAAGCTTTAAATGAATAATATGGCCAAACAGTTTATGACTATCAATATGAG	519
Qy	161	ArgLeuTrpAlaTrpGluSerTrpArgSerGluValAlaLysGlnLeuAspProLeuTyr	180
Dd	520	AGGCTCTGGGCTTGGGAAAGCTGGAGATCTGAGTCTGGCAAGCAGCTGAGGCCATTATAT	579
Qy	181	GluGluTyrValValLeuLysAsnGluMetAlaArgAlaAsnHisTyrGluAspTyrGly	200
Dd	580	GAAGGATWVGCGCTTGAAAAAAGAAATGGCAAGACCAATATATATATGAGGCTATGGG	639
Qy	201	AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyTyrTyrAspTyrSerArgGly	220
Dd	640	GATTATTTGGAAAGGAGACTATGAAAGTAAATGGGGTACATGCTATGACTACACGCCGGC	699
Qy	221	GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu	240
Dd	700	CAGTTGATTTGAAGATGTGGAAACATACCTTTGAAGACGATTTAAACCATTTATATGAAACATCTT	759
Qy	241	HisAlaTyrValAlaArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly	260
Dd	760	CATGCTTATGTGAGGCGCAAAAGTTGATTAATGCTATCTCTATATCTACGTCAATTTGA	819
Qy	261	CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpHisLeuTyrSer	280
Dd	820	TGCCCTCCCTGCTCATTTGCTTGTGGATATGTGGGGTGGATTTTGGCAAAATCTGTACTCT	879
Qy	281	LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln	300
Dd	880	TTTGACAGTTCCCTTTGGACACAAACCAACATGATGTCTATGATCAATGGTGGACCG	939
Qy	301	AlaTrpAspAlaGlnArgGliePheLysGluAlaGluLysPhePheValSerValGlyLeu	320
Dd	940	GCTCGGGATGCACAGAAATATTCAAAGAGCCGACAGATCTTGTATCTGTGGCTCTT	999
Qy	321	ProAsnMetTrpGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln	340
Dd	1000	CTATATATGACTCAAGGATTTCTGGGAAATATTCATGCTATACGGACCCAGGAAATGTTCCG	1055
Qy	341	LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet	360
Dd	1060	AAAGCAGCTGCATCCACACAGCTTGGGACCTGGGGAAGGGCCGACTTCAGAGATCCTTATG	1111
Qy	361	CysThrLysValIleThrMetAspAspPheLeuThrAlaHisGlnMetGlyHisIleGln	380
Dd	1120	TGCACAAAGGTGACAAAGGACGACTTCTACAGCTCATCTATGATGGGCGCATATCCAG	1177
Qy	381	TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgGlnLysAlaGlnGluTyrPhe	400
Dd	1180	TATATATGGCATATGCTGTCACAAACCTTTCTCTCTAAGAAATGAGACTATATCAGATTC	1233
Qy	401	HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer	420
Dd	1240	CATACAGCTGTTGGGGAATCATGTCACTTTCGCGACGCCACACTTAAGCATTTAAATATCC	1295
Qy	421	IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu	440
Dd	1300	ATTGCTCTTCTGTACCCGATTTTCAAGAAAGACAAATGAAACAGAAATTAATCTTCTGCTC	1355
Qy	441	LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg	460
Dd	1360	AAACACAGCAGTCAGATTTGTGGGACTCTCCATTTACTTACATCTTTAAGAAAGTGGAG	1411
Qy	461	TrpMetValPheLysGlyGluIleProLysAspGlnThrPheLysTyrTrpTrpIleMet	480
Dd	1420	TGGATGGCTTTAAAGGGGAAATTTCCAAAGACCAAGGATGATAAAAAAGTGGGAGATG	1477
Qy	481	LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro	500
Dd	1480	AAGGAGAGATAGTTGGGGGTGGGAAACCTGTCTCCCATGATGAAACATATCTGTACCCC	1533
Qy	501	AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTyrThrArgThrLeu	520





QY	61	Asmetsnsnsnaiaaglyaspysrtpssralapheleuylsglunserthrleu	80
Db	220	AACATGATATATGCTGGGGCAAAATGCTGCTCTTTTAAAGAACAGTCCACTTGC	279
QY	81	GlnMetYrProLeuGlnGlnIleGlnAsnLeuThrValLysLeuGlnleuGlnAlaLeu	100
Db	280	CAATGATATCCACTACAGAAATTCAGANTCTCACAGTCACACTTCAGCTGAGGCTCT	339
QY	101	GlnGlnAsnGlySerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeu	120
Db	340	CACGAAATGGGCTCTTCAGTGCCTCTCGAAGAACAAAGCAACGGTTGACACAATCTCA	399
QY	121	AsnThrMetSerThrIleYrSerThrGlnLysValCysAsnProAspAsnProGlnIu	140
Db	400	AATTCAAATGACACCACTTCACTCACTCGGAAAAGTTTGAACCCAGATATCCACAAATA	459
QY	141	CysLeuLeuLeuGlnProGlyLeuAsnGlnIleMetAlaAsnSerLeuAspYrAsnIu	160
Db	460	TGCTATTACTTGACACCAAGTTTGAAATGAATATGCGAAACAGTTTACACTACAAATGAC	519
QY	161	ArgLeuThrAlaTrpGlnSerTrpArgSerGluValGlyLysGlnLeuArgProLeuYr	180
Db	520	AGGCTCTGGGCTGGGAAACCTGGAATCTGAGAGTGGCAACACACTGGGCACTTATAT	579
QY	181	GluGluYrValValIleLysAsnGlnMetAlaArgAlaAsnHisTyrgLysAspYrGly	200
Db	560	GAAGAGTATGTGCTTGAAAATAATGAGTGGCAAGGCAAACTATTATAGAGCACTATGG	639
QY	201	AspYrTrpArgGlyAspYrGlyValAlaAsnGlyValAspGlyTyraSpYrSerArgGly	220
Db	640	GATTTTGGAGAGAGACCTTGAAGTAAATGGGGTGTGATGCTATGACTACAGCCGGCC	699
QY	221	GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuYrGlnHisLeu	240
Db	700	CAGTGTATGAAAGATGTGGAAACATACCTTTGAAGAGATTAAACCATATATGACATCTT	759
QY	241	HisAlaYrValAlaArgAlaLysLeuMetAsnAlaTyProSerTyrlIleSerProIleGly	260
Db	760	CATGCTATGTGAGCGCAAAATGTGATGAAGCTTACCTCTCAATATCAGTCCCAATTGCA	819
QY	261	CysLeuProAlaHisLeuLeuGlnLysAspMetTrpGlyArgPheTrpThrAsnLeuYrSer	280
Db	820	TGCCCTCCCTGCTATTGCTGTGGTATGTGGGGTAGATTTTGGCAATGTGACTCT	879
QY	281	LeuThrValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGln	300
Db	880	TTGACAGTCCCTTTGGACAGCAAAACCAACATATGATGTTACATGATGCAATGCTGGACAG	939
QY	301	AlaTrpAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeu	320
Db	940	GCCCTGGGATGCACAGAAATATTCAGAGAGCGCGAATGTTCTTGATCTATCTTGCTCTT	999
QY	321	ProAsnMetThrGlnGlnYrPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln	340
Db	1000	CTATATATGACTCAAGGATTTCTGGGAAAATTCATCTATACGAGACCCAGAAATTTAG	1059
QY	341	LysAlaValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMet	360
Db	1060	AAACAGACTGCCATCCCAACAGCTTGGGACCTGGGGAAGGGGACTTCAGATCTTATG	1119
QY	361	CysThrLysValThrMetAspAspPheLeuThrAlaHisGluMetGlyHisIleGln	380
Db	1120	TGCACAAAGGAGACAATGAGACACTTCCTGTACAGCTCATCATGATGAGATGGGCAATCCAG	1179
QY	381	TyrAspMetAlaTyAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGluGlyPhe	400
Db	1180	TATGATATGCGATATGCTGCACACAACCTTTCTGCTAAGAAAGAGAGCTATGAAGATTC	1239
QY	401	HisGluAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuYrSer	420
Db	1240	CATAGAGCTGTGGGGAAATCATGTCACTTTCGACGCCACACCTTAAAGATTAAATCC	1299

QY	421	ILISLVLEU	LeuSerProAspPheGlnGluAsnGluThrGluLeu	440
Db	1300	ATTGGCTTCTGTCA	CCCGATTTCAGAAAGCAATGAACAAATTAACCTTCCTGCTC	1355
QY	441	LysGlnAlaLeuThrIleVal	IGlyThrLeuProPheThrTyrMetLeuGluIubStrParG	460
Db	1360	AAACAGACAC	TCACAGATTGGTTGGGACTCTGCATTTACTTACATGTTAGACAAAGTGGAG	1419
QY	461	TrpMetValPheIysGlyGluIlePro	LysAspGlnTrpMetLysLysTrpTrpLeuMet	480
Db	1420	TGGATGGCTTTAA	AGGGGAAATCTCCAAAGACCACTGGATGTAAAAAGGTGGGAGATG	1479
QY	481	LysArgGluIleVal	IGlyValGluProValProHisAspGluThrTyrCysAspPro	500
Db	1480	AAGCAGAGATAGT	GGGGGTGGTGGACCTGTGCCCATGATGAACACTGTGATCCCC	1538
QY	501	AlaSerLeuPheHisValSerAsn	AspTyrSerPheIleArgTyrThrThrArgThrLeu	520
Db	1540	GCATCTCGTTCAT	GGTTTTCGATGATTACTCATTCATTCGATTAACACAAAGGACCCCT	1599
QY	521	TyrGlnPheGlnPheGlnGluAlaLeu	CysGlnAlaAlaLysHisGluIleProLeuHis	540
Db	1600	TACCAATTCAGATTCA	AGAAAGCACTTTGTCAAGCAGCTAAACATGAAAGGCCCTCGAC	1659
QY	541	LysCysAspIleSerAsnSerThrGlu	AlaGlyGluLysLeuPhe-----	555
Db	1660	AAATGTGACATCTCA	AACTCTACAGAAAGCTGGACAGAACTGTT-GTAAGAAATACCTCA	1718
QY	555	-----	-----	555
Db	1719	AAATGTTCAACCTCC	AGATCTCAGATTACTCATTTCCATTCGCTAGCTTGTATTTGG	1778
QY	555	-----	-----	555
Db	1779	ATTTCCTTTGTTCT	AAAAAGAAATTTTATGGCCTCAAAATGTCCTCATTTACAAACAAA	1838
QY	555	-----	-----	555
Db	1839	CATTTAATTTGTGGT	CAGACAGCAACTAGACATCAACAATTTGGGTGGGCCACCTCTT	1898
QY	555	-----	-----	555
Db	1899	TTTCTCCCATCAT	ATACACAGCCCTCTCTCCGTGGATATTGGAGAGAAAGACGGTTAG	1958
QY	555	-----	-----	555
Db	1959	GGTGAATATATCT	GTTAATATGCAATCTTCTTATCTGCCAGAACAAATTTAGCCAA	2018
QY	555	-----	-----	555
Db	2019	GTCAAAGAGAA	AAACCATAGATCATAGATGAATATATATGATCATCTGAAACCCCTCAA	2078
QY	556	-----	-----AsnMetLeuArgLeuGlyLysSerGluPro	565
Db	2079	AAGGCCCTGAAC	CCCCCTTTTGTGTAGCAAAATATGCTGAGGCTTGGAAAATTCAGAACCC	2138
QY	566	TrpThrIleuAlaLeuGluAsnVal	IleValGlyAlaLysAsnMetAsnValArgProLeuLeu	585
Db	2139	TGGACCTTACGAT	TGGAAAAATGTTGTAGAGC-AGAACATGAAATGTAAGGCCACATGCTC	2197
QY	586	AsnTyrPheGluIleProLeuPheThr	TrpLeuLysAspGlnAsnLysAsnSerPheValGly	605
Db	2198	AACTACTTTTGG	CCCCATTATTACTGGCTCAAAAGACAGAACMAAGAAATCTTTTGGGGA	2255
QY	606	TrpSerThrAspTrpSerProTyr	AlaAsp-GlnSerIleLysValArgIleSerLeuLys	625
Db	2258	TGGATATACGAC	TGGAGTCCATATGCGACCAAGCATCAAGATGAAGTGAAGATTAAGCTTAAA	2311
QY	625	SerAlaLeuGlyAspLysAlaTyrGlu	TrpAsnAspAsnGluMetTyrLeuPheArgSe	645
Db	2318	ATCAGACTCTTGG	AGATAAGCATATGAAATGGAACGAAATGAATATGCTGTTCGATAC	2377
QY	645	rserValAlaTyrAlaMet	ArgGlnTyrPheLeuLysValLysAsnGluMetIleLeuPhe	665

```

|||||
Db 2378 ATCTGTGATATGCTATGAGGACGACTTTTAAAGAAAAATCAGATGATCTTTT 2437
Qy 665 eelglucluaspyvalaigvalaAlaasnleuysProargileSerpheasnphheva 685
Db 2438 TGGGAGGAGAGATGCGAGTGGCTAAATTGAAACCAAGAAATCTCTTAATTCTTGT 2497
Qy 685 lthrlapPolysasnvalSerAspIleleProarghrGluValGluValAlaIleAr 705
Db 2498 CACTGCACCTAAAAATGCTCTGATATATCTTCTTAAGAACTGAAGTTGAAGGCCATCAG 2557
Qy 705 gneSerArgSerArgIleasnPalapheArgleuasnAspaSnSerleuGluPhele 725
Db 2558 GATGTCGCGAGCCGATCAATGATGCTTCGCTGATGATGACACACGCTAGATTTCT 2617
Qy 725 uGlyIleGlnProthrleuGlyProProAsnGlnProValSerIletrPleuIleVa 745
Db 2618 GGGGATACAGCCACACCTTGACCTTACACAGCCCTGTTCCATATGCGTATGTTGT 2677
Qy 745 lPheGlyValValMetGlyValIleValValGlyIleValIleleuIlePheThrGlyI 765
Db 2678 TTTTGATTTGATGGAGAGATGATGCTTGGCATTTGTCATCCTGATCTTCACCTGGAT 2737
Qy 765 eArgAspArgIleuysAsnLysAlaArgSerGlyLysAsnProtyrAlaSerIleas 785
Db 2738 CAGAGATCGAAGAAAGAAATAAAGCAAGAGTGAGAAATCCTATGCTCCATCGA 2797
Qy 785 pIleSerIysGlyLysAsnAsnProGlyPheGlnAsnThrAspAspValGlnThrSerph 805
Db 2798 TATTGCAAGAGAGAAATATATCCAGATTCCTCAAAACACTGATGATGTCAGACCTCTT 2857
Qy 805 e 805
Db 2858 T 2858

RESULT 15
US-09-969-384-2
; Sequence 2, Application US/09969384
; Publication No. US20020192749A1
; GENERAL INFORMATION:
; APPLICANT: Moore, et al.
; TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PTO55PI
; CURRENT APPLICATION NUMBER: US/09/969,384
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: PCT/US01/10542
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/236,384
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/194,118
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2920
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1707)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2702)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2749)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2757)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2788)
; OTHER INFORMATION: n equals a,t,g, or c

```

```

; NAME/KEY: SITE
; LOCATION: (2789)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2819)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2835)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2856)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-969-384-2

Alignment Scores:
Pred. No.: 0
Score: 4061.00
Percent Similarity: 99.35%
Best Local Similarity: 99.22%
Query Match: 94.64%
DB: 9
Gaps: 0

US-09-978-385-2 (1-805) x US-09-969-384-2 (1-2920)
Qy 3 SerSerSertrPleuLeuSerleuValAlaValThrAlaAlaGlnSerThrIleGlu 22
Db 35 AGCTCTCCGCGCTCTTCCACCTTGTGCTGTAACGCTGCTCAGTCCACCATTTAG 94
Qy 23 GluGlnAlaLysThrPheLeu-AspLysPheAsnHisGluAlaGluAspLeuPheTrgI 42
Db 95 GAACAGGCCAAGACATTTTGGGACCAAGTTTAAACCGAAGCCGAAGACCTGTTCTATCA 154
Qy 42 nSerSerleuAlaSerTrpAsnTrpAsnThrAsnIleThrGluGluAsnValGlnAsnke 62
Db 155 AAGTTCACCTTGTCTTGGATATATTAACACCAATATTACTGAAGAGATGTCACAAACAT 214
Qy 62 lAsnAsnAlaGlyAspLysTrpSerAlaPheLeuysGluGlnSerThrleuAlaGlnme 82
Db 215 GAATTAATGCTGGGGACAAATAGTCTGCTTTTAAAGGAACAGTCCACACTTGCCAAAT 274
Qy 82 tTryProLeuGlnGluIleGlnAsnleuThrValLysleuGlnLeuGlnAlaLeuGlnI 102
Db 275 GATTCACACTAACAAGAAATTCAGATCTCAGATGCAAGCTTACGCTGACAGCTCTTCAGCA 334
Qy 102 nAsnGlySerSerValleuSerGlyLysAspLysSerIysArGleuAsnThrIleLeuAsnTh 122
Db 335 AAATGGGCTTTCAGTGCCTCTCAGAAAGCAAGCAAAACGGTGAACACAAATCTAAATAC 394
Qy 122 rMetSerThrIleTrySerThrGlyLysValCysAsnProAspAsnProGlnGluCysLe 142
Db 395 AATGACACACATCTACAGTACTGGAAGAGTTTGTAAACCCAGATTAATCCACAGAAATGCTT 454
Qy 142 uLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerleuAspTryAsnGluArgLe 162
Db 455 ATTACTTGAACACAGGTTTGAATGAATATATGCAAAACAGTTTAGACTCAATGAGAGGCT 514
Qy 162 uTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuAArgProLeuTryGluG 182
Db 515 CTGGGCTTGGGAAAGCTGGAGATCTGAGGCTGGCAAGCAGCGTGAACCATTAATAGAGA 574
Qy 182 uTryValValleuLysAsnGluMetAlaArgAlaAsnHisTryGluAspTryGlyAspTry 202
Db 575 GATGAGGCTGTAAGAAATGAGATGGCAAGATGCAATATCTTATGAGACTATGGGATTA 634
Qy 202 rTrpArgGlyAspTryGluValAsnGlyValAspGlyTryAspTrySerArgGlyGlnLe 222
Db 635 TTGGAGAGGAGACTATGAGTAATGAGGATGATGCTATGACTACACCCGCGCAAGT 694
Qy 222 uIleGluAspValGluHisThrPheGluGluIleLysProLeuTryGluHisAla 242
Db 695 GATGAGAGTGTGGAACATCTTTGAAGAGATTAAACCATTAATATGAACATCTTCATCAG 754
Qy 242 atyValAlaArgAlaLysLeuMetAsnAlaLysProSerTryIleSerProIleGlyCysLe 262

```

```

|||||
Db 755 CTATGTAGGCGCAAGTTGATGATGCTTCTCTTCTATATGATCCATTGGATGCTT 814
QY 262 uPrOAlAHISleuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTyrSerLeuTh 282
Db 815 CCCGCGCATATTTGGTGTATGTGGGTAGATTTTGGACAATAATGTGACMSTTTGAC 874
QY 282 rValProheGlyGlnYsProAsnIleAspValThrAspAlaMetValAspGlnAlaTr 302
Db 875 AGTTCCCTTTGGACAGAAACCAACATAGATGTTACTGATGCAATGGTGGACCAAGRCTG 934
QY 302 pAspAlaGlnArgIlePheLysGlnAlaGlnLysPhePheValSerValIleLeuProAs 322
Db 935 GGATGCGACAGAAATATTCAGAGAGCGCGAAGTTCTTTGTATCTTGGTCTCTTAA 994
QY 322 nMetThrGlnIlyPheTrpGlnAsnSerMetLeuThrAspProGlyAsnValGlnLysAl 342
Db 995 TATGACTCAAGATTTGGGAAATTCATGCTAACGACCCAGGAATGTTCCAGAAAGC 1054
QY 342 aValCysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMetCysTh 362
Db 1055 AGTGGCCATCCCAAGCTTGGAGCTGGGGAAGGCGACTTCAGATCTTATGTGCAC 1114
QY 362 rLysValThrMetAspAspPheLeuThrAlaHisHisGlnMetGlyHisIleGlnTyrAs 382
Db 1115 AAAGGTGCATGAGCAGCTTCGACAGCTCATGAGATGGGCGATATCCAGTATGA 1174
QY 382 pMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlnGlyPheHisG 402
Db 1175 TATGGCATATGCTGCACAACTTTTCTGTAAAGAAATGGACCTATGAAAGATTCATGA 1234
QY 402 uAlaValGlyGlnIleMetSerLeuSerAlaIleThrProLysHisLeuLysSerIleG 422
Db 1235 AGCTGTTGGGAATCATGCTACTTCTGCAGCCACACTTAAAGATTTTAAATCATTTGG 1294
QY 422 yLeuLeuSerProAspPheGlnGlnLysAspAsnGlnuThrGlnIleAsnPheLeuLysG 442
Db 1295 TCTTGTGACCCGATTTTCAAGAAAGACATGAAGAAATAAAGCTTCGTCACAAACA 1354
QY 442 nAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnLysTrpArgTyrPme 462
Db 1355 AGCACTCAGATTTGGGAGCTGCTCCATTACTTACATGTTAGAGAAATGGAGGTGAT 1414
QY 462 tValPheLysGlyGlnIleProLysAspGlnTrpMetLysTrpTrpGlnMetLysArg 482
Db 1415 GGTCTTTAAAGGGAATTTCCAAAGACAGTGGATGAAAAAGTGTGGAGATGAAGAGC 1474
QY 482 gGlnIleValGlyValGlnProValProHisAspGlnuThrTyrCysAspProAlaSe 502
Db 1475 AGAGATAGTGGGTGGTGGAGAACTGCCCCCATGATGAAACATCTGTGACCCCGCATC 1534
QY 502 rLeuPheHisValSerAsnAspTyrSerPheIleArgTyrTrpThrArgThrIleuTyrG 522
Db 1535 TCTGTTCATGTTTCTAATGATTTACTCATTTCTGATATTTACACAAAGGACCTTTACCA 1594
QY 522 nPheGlnPheGlnGlnAlaLeuGlnLysGlnAlaAlaLysHisGlnGlyProLeuHisLysCy 542
Db 1595 ATTCCAGTTTCAAGAGCACTTTGTCAAGCAAGCTAAACATGAAGCCCTTCGACAAATG 1654
QY 542 sAspIleSerAsnSerThrGlnAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLy 562
Db 1655 TGACATCTCAAACTACAGAGAGCTGACAGAAACTGTTCAATATGCTGAGAGNTTGA 1714
QY 562 sSerGlnProTrpThrIleAlaLeuGlnLysValValGlyAlaLysAsnMetAsnValAr 582
Db 1715 ATTCAGAACTTGACCTTAGCATGGAATGTTGTAGAGCAAGAAACATGATGTAAAG 1774
QY 582 gProLeuLeuAsnTyrPheGlnProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSe 602
Db 1775 GCCACGCTCAACTACTTGGACCTTATTTACCTGGGCGAAAGACCAAGAAATATTC 1834
QY 602 rPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleLysValArgI 622
|||||

```

```

Db 1835 TTTTGGGATGAGTACGAGTACGAGTCCATATGACAGCAACCAAAAGTGGAGAT 1894
QY 622 aSerLeuLysSerAlaLeuGlyAspLysAlaIleTrpGlnTrpAsnAspAsnGlnMetTyrLe 642
Db 1895 AAGCTTAAATATGAGCTTTGGAGATTAAGCATATGATGAAGCAATGAAATGTACTT 1954
QY 642 uPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsnGlnMe 662
Db 1955 GTTCCGATCATCTGTTGCAATATGATGAGCGCACTTTTAAAGTAAATAATCAGAT 2014
QY 662 tIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProAlaGlyIleSerPheAs 682
Db 2015 GATTCCTTTTGGGAGAGATGTCGAGTGGCTAATTTGAACCAAAATCTCTTTAA 2074
QY 682 nPhePheValThrAlaProLysAsnValSerAspIleIleProArgTrpGlnValGlnLys 702
Db 2075 TTTCTTTGCTACTGCACTTAAATATGTCTATATCATCTTCAACTGAAGTTGAAA 2134
QY 702 sAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLe 722
Db 2135 GGCCATCAGAGATGTCGGGAGCCGTATCAATGATGCTTCCGTCGAATGACAGACGCT 2194
QY 722 uGlnPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSerIleTr 742
Db 2195 AGAGTTTCTGGGATACACCAACACTTGACCTCTTAAACAGCCCTGTTCCATATG 2254
QY 742 pIleuIleValPheGlyValValMetGlyValIleValValGlyIleValIleLeuIlePhe 762
Db 2255 GCTGATTTGTTTGGAGTTGTATGGAGTGTATGAGTGTGGCATTTGTCATCTGATCT 2314
QY 762 eThrGlyIleArgAspArgLysLys 770
Db 2315 CACTGGATCAGAGATCGAAGAG 2339

RESULT 16
US-09-969-384-12
; Sequence 12, Application US/09969384
; Publication No. US20020192749A1
; GENERAL INFORMATION:
; APPLICANT: Moore, et al.
; TITLE OF INVENTION: Human Gene Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT055PI
; CURRENT APPLICATION NUMBER: US/09/969,384
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: PCT/US01/10542
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/236,384
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/194,118
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 2911
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1705)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (2847)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-969-384-12

Alignment Scores:
Pred. NO.: 0
Score: 4013.00
Percent Similarity: 99.35%
Best Local Similarity: 99.22%
Query Match: 93.52%
Gaps: 0
Length: 2911
Matches: 763
Conservative: 1
Mismatches: 4
Indels: 3
Gaps: 0

```

us-09-978-385-2 (1-805) x us-09-969-384-12 (1-2911)

QY 3 SerSerSerThrLeuLeuLeuSerLeuValAlaValThrAlaAlaGlnSerThrIleGlu 22  
 Db 35 AGCTCTTCGGGCTCCCTTCACGCTTGTGCTGAACCTGCTCCTCAGTCACCAATGAG 94  
 QY 23 GluGlnAlaIleThrPheLeu-AspLysPheAsnHisGluAlaGluAspLeuPheArgI 42  
 Db 95 GAACAGGCCAAGACATTTTGGGACACAAGTTTAAACCAGAGCCGAGACCTGTCTATCA 154  
 QY 42 nSerSerLeuAlaSerTrpAsnTrpAsnThrAsnIleThrGluGluAsnValGlnAsnMe 62  
 Db 155 AAGTTCACCTGCTCTTGGAAATTATTAACACCATATTAGCAAGAGAAATGCCAATAACAT 214  
 QY 62 LAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAlaGlnMe 82  
 Db 215 GAATTAATGCTGGGACAAATGCTCCTTTTAAAGAACACTCCACACTGCCCCAAAT 274  
 QY 82 tTyProLeuGlnGluIleGlnAsnLeuThrValLysLeuGlnLeuGlnAlaLeuGlnGlu 102  
 Db 275 GATTCACACTACAGAAATTCAGAAATCTCACAGTCAAGCTTCAGCTCAGGCTTTCAGCA 334  
 QY 102 nAsnGlySerSerValLeuSerGluAspLysSerLysArgLeuAsnThrIleLeuAsnTh 122  
 Db 335 AAATGGGTCTTCAGTCTCTCAGAAAGACAAAGCAACGAGTTGAACCAATTCTAATATAC 394  
 QY 122 rMetSerThrIleTySerThrGlyLysValCysAsnProAspAsnProGlnGluCysLe 142  
 Db 395 AATGACACACCATTCACAGTACAGTACGAAATTTGTATACCCAGATATCCACAAGAAATGCT 454  
 QY 142 uLeuLeuGluProGlyLeuAsnGluIleMetAlaAsnSerLeuAspTyArgLysGlu 162  
 Db 455 ATTACTTGAAACGAGTTTGAATGAATATGCAAAACAGTTTAAAGTACAGAGGCT 514  
 QY 162 uTrpAlaTrpGluSerTrpArgSerGluValGlyLysGlnLeuArgProLeuTyGluGlu 182  
 Db 515 CTGGGCTTGGGAAAGCTGAGATCTGAGGCTGGCAAGCAGTGGCCCTTTATATGAGAA 574  
 QY 182 uTyValValLeuLysAsnGluMetAlaArgAlaAsnHisTyGluAspTyArgLysPty 202  
 Db 575 GATATGGCTTGAATAAATGAGATGCGACAGCAACCAATCTTTATGAGAGCTATGGGATTA 634  
 QY 202 rTrpTrpArgLysAspTyArgLysValAsnGlyValAspGlyTyArgPtySerArgLysGlu 222  
 Db 635 TTGGAGAGGAGACTATGAGTAATGAGGCTAGATGGCTATGACTACAGCCGGCCAGT 694  
 QY 222 uIleGluAspValGluHisThrPheGlnGluIleLysProLeuTyGluHisLeuHisAl 242  
 Db 695 GATTCAGATGTGGACATACCTTTGAGAGAGATTAACCCATTAATGAAACATCTTCATGC 754  
 QY 242 aTyValAlaArgAlaLysLeuMetAsnAlaTyProSerTyIleSerProIleGlyCysLe 262  
 Db 755 CTATGAGAGCCAAAGTTGATGAATGCCATCTCCATATCAATCCCAATGGATGCC 814  
 QY 262 uProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpThrAsnLeuTySerLeuTh 282  
 Db 815 CCCTGCTCATTTGCTGGATATGTGGGCTAGATTTGGACAATCTGTACTCTTGAC 874  
 QY 282 rValProPheGlyGlnLysProAsnIleAspValThrAspAlaMetValAspGlnAlaTr 302  
 Db 875 ACTTCCCTTTGGACAGAAACCAACATAGATGTACTAGTGAAGAGTGGACAGCCGCTG 934  
 QY 302 pAspAlaGlnArgIlePheLysGluAlaGluLysPhePheValSerValGlyLeuProAs 322  
 Db 935 GGATGCACACGAGAAATATTCAGAGAGCGGAGAGTCTT-GRATCTGTGGCTTCTCA 993  
 QY 322 nMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGlnLysAl 342  
 Db 994 TATGACTCAAGATCTCTGGAAATTCATGCTTAACGAGACCCAGAGAAATGTTCAAGAAAGC 1053  
 QY 342 aValLysHisProThrAlaTrpAspLeuGlyLysGlyAspPheArgIleLeuMetCysTh 362  
 Db 1054 AGTCTGCATCCCAAGCTTGGACCTGGGGAGGGCGACTTCAGAGATCTTATGTGAC 1113

QY 362 rLysValThrMetAspPheLeuThrAlaHisHisGluMetGlnHisIleGlnTyArg 382  
 Db 1114 AAAGTGCACAAATGACGACTTCCTGACAGCTCATATGAGATGGGCAATATCCAGTAA 1173  
 QY 382 pMetAlaTyAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyPheHisGlu 402  
 Db 1174 TATGGCATATGCTGCACAACCTTTCTGTAAGAAATGAGCTAATGAAAGATTCATCA 1233  
 QY 402 uAlaValGlyGluIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSerIleG 422  
 Db 1234 AGCTGTGGGCAATCATGCTACCTTTCTGCAGCCACACCTTAAGCATTTTAAATCCATTGG 1293  
 QY 422 yLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLysGlu 442  
 Db 1294 TCTTCTGCACCCGATTTTCAAGAAAGACATGAACAAATTAACCTTCGCTCAAAACA 1353  
 QY 442 nAlaLeuThrIleValGlyThrLeuProPheThrTyMetLeuGluLysTrpArgTrpMe 462  
 Db 1354 AGCACTCAGATGTTGGGAGCTCTGCATTTACATTAAGATGAAGAGTGGAGTGGAT 1413  
 QY 462 tValPheLysGlyGluIleProLysAspGlnTrpMetLysTrpTrpGluMetLysArg 482  
 Db 1414 GGTCTTTAAAGGGGAAATTCACCAAGACACAGTGAAGAAAGTGGGAGATGAAGCG 1473  
 QY 482 gGluIleValGlyValGluProValProHisAspGluThrTyArgCysAspProAlaSe 502  
 Db 1474 AGAGATAGTGGGGTGGGAACTGTGCCCATGATGAACATATCTGTGACCCCGCATC 1533  
 QY 502 rLeuPheHisValSerAsnAspTySerPheIleArgTyTrpThrArgThrLeuTyArg 522  
 Db 1534 TCTGTTCATGTTCTTAATGATTAATCTCATTTATTCGATTAACAAAGAGACCTTTACCA 1593  
 QY 522 nPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluLysProLeuHisLysCy 542  
 Db 1594 ATTCAGATTTCAAGAGACATTTGTCAAGCAGTAAACATGAAGCCCTCGCAAAATG 1653  
 QY 542 sAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeuGlyLy 562  
 Db 1654 TGACATCTGC-AACCTCAGAGAGCTGGACAAACCTGTCAATATGCTGAGGNTGGAA 1712  
 QY 562 sSerGluProTrpThrLeuAlaLeuGluAsnValValGlyAlaLysAsnMetAsnValAr 582  
 Db 1713 ATCAAGAACCCGAGACCTCAGCATTTGGAAATGTTGTAGGACCAAGACATGAATGTAAG 1772  
 QY 582 gProLeuAsnAsnTyPheGluProLeuPheThrTrpLeuLysAspGlnAsnLysAsnSe 602  
 Db 1773 GCCACTGCTCAACTACTTTGAGCCCTTATTTACTGCTGCGTGAAGACCAAGAAATTC 1832  
 QY 602 rPheValGlyTrpSerThrAspTrpSerProTyAlaAspGlnSerIleLysValArgI 622  
 Db 1833 TTTGTGGAGATGAGTACCGAGCTGAGATCCATATGAGACCAAGCAATCAAGATGAGAT 1892  
 QY 622 eSerLeuLysSerAlaLeuGlyAspLysAlaTyArgIleTrpAsnAspAsnGluMetTyLe 642  
 Db 1893 AAGGCTAAATACACTCTTGGAGATTAAGCATATGATGAGACAGACATGAATGTAACCT 1952  
 QY 642 uPheArgSerSerValAlaTyAlaMetArgGlnTyPheLeuLysValLysAsnGlnMe 662  
 Db 1953 GTTCGATCATCTGTTGCATATGCTATGAGGCGAGTACTTTTAAAGTAAATAAATCAGAT 2012  
 QY 662 tIleLeuPheGlyGlnGluAspValArgValAlaAsnLeuLysProArgIleSerPheAs 682  
 Db 2013 GATTCCTTTTGGGAGGAGATGTGGCAGTGAATTTGAACCAAGAAATTCCTTAA 2072  
 QY 682 nPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluValGly 702  
 Db 2073 TTTCTTTGTCACGACCTAAATATGCTGTATCATCTTCATGAACATGAATTAATAA 2132  
 QY 702 sAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsnSerLe 722  
 Db 2133 GGCATCAGAGATGTCGCGGAGCCGTATCAATGATGCTTCCGTGAATGAGACAGACGCT 2192











```

? PRIOR FILING DATE: 2000-05-03
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO: 3
? LENGTH: 2415
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: This degenerate sequence encodes the amino acid
? OTHER INFORMATION: sequence of SEQ ID NO:2.
? NAME/KEY: misc_feature
? LOCATION: (1)..(2415)
? OTHER INFORMATION: n = A,T,C or G
? OS-09-978-385-3

```

### Alignment Scores:

Pred. No.:	0	Length:	2415
Score:	3509.00	Matches:	644
Percent Similarity:	80.00%	Conservative:	0
Best Local Similarity:	80.00%	Mismatches:	161
Query Match:	81.78%	Indels:	0
DB:	9	Gaps:	0

US-09-978-385-2 (1-805) x US-09-978-385-3 (1-2415)

[illegible]

QY	241	HSLSATATYVALARGVALISLEULEMELFASALATYPRoserYrTLIeserProIleGly	260
Db	721	CAYGCNTATGTMMGNCMAAYTNATPAAGCCNATCCMSNTNATVTHMSNCCNATHGN	780
QY	261	CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgpHeTrpHisLeuIysSer	280
Db	781	TCYTYNCCNCCNCAAYTNATYTNMGNGAATATGTGGGMMGNNTTYGGACMAAYTTNATYMSN	840
QY	281	LeuThrValProPheGlyGlyValysProAsnIleAspValThrAspAlaMetValAspGln	300
Db	841	YTNACNGNCCNTTYGGNCACAAACRCMAAATGAAGTNNACNCAVCNATGGTNCACAR	900
QY	301	AlaTrpAspAlaGlnAcGlyLePheIysGluAlaGlnIlySphetheValSerValGlyLeu	320
Db	901	GCNTGGGAYGCNCARMGNATHTTAAAGACGCNGARARATTTTGTGTMNSGNTGCGNTTN	960
QY	321	ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln	340
Db	961	CCNAAYATGACNCARCAGNTTYTGGAGAAAYWNAATGTYNNACNGACVCCNGNAAYTCNCAR	1020
QY	341	LysAlaValCysHisProThrAlaTrpAspLeuGlyIlySgIlyAspPheArgIleLeuMet	360
Db	1021	AARGCNGNTYCAVCCNACNGCTGGGAYTGTGNGNAAGCGCATYTTGNTAHTYNTAGT	1080
QY	361	CysThrIysValThrMetAspAspPheLeuThrAlaHisIleGlnMetCylHisIleGln	380
Db	1081	TGYACNARAGTNACNATGAGATYTYTNACGNCACATCANGARTGGGCNATYTHCAR	1140
QY	381	TyrAspMetAlaTrpAlaAlaGlnProPheLeuLeuArgAsnGlyAlaAsnGlyIlyPhe	400
Db	1141	TAYGATGAGCMTAYGCNGNCARCCTTYTYTNMGNAAGAGCNCMAAYCARCGNTTY	1200
QY	401	HisGluAlaValGlyIlyIleMetSerLeuSerAlaAlaThrProIlyHisLeuIysSer	420
Db	1201	CATGARGCNGTNGNGARATHTATGMSNTYTNMSNGCNCACCCNMAARCATYTNARWSN	1260
QY	421	IleGlyLeuLeuSerProAspPheGlnGlyAspAsnGlnThrGluIleAsnPheLeuLeu	440
Db	1261	ATHGAGTNTNTNMSCNCGATYTYTCARGARAGAAAGACAGACGARTHAATYTYTNTN	1320
QY	441	LysGlnAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGlnIlyStrParG	460
Db	1321	AARCARCCTNTMCNATHTGNGNACNNTYNTACNTAATATYTGARARARAGGNGN	1380
QY	461	TrpMetValPheIysGlyGlyIleProIlyAspGlnTrpMetIlySgIlyStrTrpGluMet	480
Db	1381	TGATGTGNTTYYAARGNGARATHTCCMAARGATCATGATGAATAAATGTGGGARATG	1440
QY	481	LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro	500
Db	1441	AARGNCAARATHTGNGNGTNGNARCNCNGNCCNCAAGAAGACNATATGAYACCN	1500
QY	501	AlaSerLeuPheHisValSerIleAsnArgYrSerPheIleArgTyrTyrThrArgThrLeu	520
Db	1501	GCNMSNTNTTTCAYGTMNSMAAYGATYATWNTTYATHMGNTATATACMMGNACAYTN	1560
QY	521	TyrGlnPheGlnPheGlnGlyAlaLeuCysGlnAlaAlaIlyHisIlySgIlyProLeuHis	540
Db	1561	TATACARTTCARTTYCARARGCNGNTNTTYCARCGNCCNMAARCAAGARGGCCNYTTCAY	1620
QY	541	LysCysAspIleSerAsnSerThrGluAlaGlyGlnIlySleuPheAsnMetLeuArgLeu	560
Db	1621	AARGTAYATHTMSNNAAYMSNMCNARGCNGSNCARARATYTTTAAATGTMMGNNTN	1680
QY	561	GlyIlySerGluProTrpThrLeuAlaLeuGlnAsnValValGlyAlaIlyAsnMetAsn	580
Db	1681	GGNAAMMSNARCNCNCGACNTYNGCNATYNGARAAAYGTNGTNGCNCMAAATAATGAAY	1740
QY	581	ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuIlySAspGlnAsnIlyS	600
Db	1741	GTMNGNCNTYTNNAATYATYTYTCARCCNTYNTTYACNTGTYTNAAGATCAAAATAAR	1800
QY	601	AsnSerPheValGlyTrpSerThrAspTrpSerProTyrAlaAspGlnSerIleIysVal	620

```

Db 1801 AAYMSNTTYGTGNTGMSNACNGATYGGMSNCCNTAYGCNGAYCARMSNATRAKGTN 1860
Qy 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspGluMet 640
Db 1861 MGNATHMSNTYNAAMWSNGCNTNGNAGAYARAGCNTAYGARTRGGAAYGAAVAGARATG 1920
Qy 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgIleTyrPheLeuLysValLysAsn 660
Db 1921 TAYNTNTTYMGMSNMSNGTNGCNTAYGCNATGMCNCRATYTYTYYTNAAGTNAARAAY 1980
Qy 661 GluMetIleLeuPheGluGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680
Db 1981 CARATGATHTYTTTGGNGARAGARCATGTMGNGTNGCNMAAYTNAACCMGNATHMSN 2040
Qy 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgTrpGluVal 700
Db 2041 TTYAAYTYYTGTNACGNCNCNNAARAAYGTNMGNGAYATHATHTCCNMGNAACNGARGTN 2100
Qy 701 GluLysAlaIleArgMetSerArgSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
Db 2101 GARARAGCAATHTMGATGMSNMGNSNMGNATHAAYGAYGCTTYMGNTYNAAGAYAAAY 2160
Qy 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
Db 2161 WSNYNTNGATYYTNGNATHCARCNCNNTNGNCCNCAATCARCNCNCCNNTMSN 2220
Qy 741 IleTrpLeuIleValPheGlyValAlaMetGlyValIleValValGlyIleValIleLeu 760
Db 2221 ATHTGYNATHTGTTTGGTNGTNGTNTATGGGNTNATHTGNTGNAATGTTNATHTTN 2280
Qy 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
Db 2281 AHTHTTACGNAATHTMGATGMSNMGNAARAARAARAARAARGCMGMSWNGMGARARATCN 2340
Qy 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGluAsnThrAspAsp 800
Db 2341 TAYGCMWSNATHTGATHTMSNARAGNGARAYAYCCNGGNTTTCARAAVACNGAYGAY 2400
Qy 801 ValGlnThrSerPhe 805
Db 2401 GTNCARACACMSNTTY 2415

RESULT 20
US-09-978-385-7
; Sequence 7, Application US/09978385
; Patent No. US20020177211A1
; GENERAL INFORMATION:
; APPLICANT: Piddington, Christopher S.
; APPLICANT: Petrie, Charles
; APPLICANT: Shoemaker, Kimberly E.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: ZACE2: A HUMAN METALLOENZYME
; FILE REFERENCE: 99-24CI
; CURRENT APPLICATION NUMBER: US/09/978, 385
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/133,952
; PRIOR FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 60/151,181
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: 09/563,516
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2415
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: This degenerate sequence encodes the amino acid
; NAME/KEY: misc_feature
; LOCATION: (1)...(2415)

```

```

; OTHER INFORMATION: n = A,T,C or G
US-09-978-385-7
Alignment Scores:
Pred. No.: 0
Score: 2904.00
Percent Similarity: 71.18%
Best Local Similarity: 65.47%
Query Match: 67.68%
DB: 9
Gaps: 0

US-09-978-385-2 (1-805) x US-09-978-385-7 (1-2415)
Qy 1 MetSerSerSerSerThrPheLeuLeuSerLeuValAlaValThrAlaGlnSerThr 20
Db 1 ATGMSNMSNMSNMSNTGGTYNTNTYNTNWSNTYNTGTCGNTNACNCCNCAKMSNTYN 60
Qy 21 IleGluGluGluAlaLysThrPheLeuAspLysPheAsnIleGluAlaGluAspLeuPhe 40
Db 61 ACNGARARARAAGCNARACNTTYYTNAAYATTAAYAACARCARAGCNGARAGAYTNMSN 120
Qy 41 TyrGlnSerSerLeuAlaSerTrpAsnTyrAsnThrAsnIleThrGluLysAsnValGln 60
Db 121 TAYCARMSNMSNTYNTGCMNSNTGAAATAYAAACNAAYATHTACNARARARAGCGCAR 180
Qy 61 AsnMetAsnAsnAlaGlyAspLysTrpSerAlaPheLeuLysGluGlnSerThrLeuAla 80
Db 181 AARATGMSNARCGNCGNCCNAAARTGMSNCGTYYTATGARARARARARARARARARAR 240
Qy 81 GluMetTyrProLeuGlnGluIleGlnAsnLeuThrValLysLeuGluLeuGluAlaLeu 100
Db 241 CARMSNTTYMSNTYNTCARATHTCARACNCCNATHTAARAGNARARNTNARCNRYTN 300
Qy 101 GlnGlnAsnGlySerSerValLeuSerGlyAspLysSerLysArgLeuAsnThrIleLeu 120
Db 301 CARARMSNCGMSNMSNCGNTYNTMSNCGATTAARAARAARARARARARARARARARAR 360
Qy 121 AsnThrSerThrIleTyrSerThrGlyLysValCysAsnProAspAsnProGlnGlu 140
Db 361 AAYACNATGMSNACNATHTAYWMSNACNGNARAGTNTGVAAYCCNABARAYCCNARCAR 420
Qy 141 CysLeuLeuLeuGluProGlyLysAsnGluIleMetAlaAsnSerLeuAspTyrAsnGlu 160
Db 421 TGYTNTNTYNTNGARCGNCGNTYNTGATGATHTATGCGNACMSNACNGATYTAAYWSN 480
Qy 161 ArgLeuTrpAlaTrpGluSerTrpArgSerGlyValGlyLysGlnLeuArgProLeuTyr 180
Db 481 MGNATNTGGGCGNTGGGARGGNTGCMGNCNGARAGTNGNARARARARARARARARARAR 540
Qy 181 GluGluTyrValAlaLeuLysAsnGluMetAlaArgAlaAsnIleTyrGluAspTyrGly 200
Db 541 GARGARATYAGTNGTNTYNTNARAAYAGARATGCMNCGNCAAYTAAYAAVAGAYTAGGN 600
Qy 201 AspTyrTrpArgGlyAspTyrGluValAsnGlyValAspGlyLysAspTyrSerArgGly 220
Db 601 GAYTAYTGGMGNGNGAYTAYGARGCGNARGCMGNCAGYNTAAYTAAYAAVAGMNAAY 660
Qy 221 GlnLeuIleGluAspValGluHisThrPheGluGluIleLysProLeuTyrGluHisLeu 240
Db 661 CARITNATHGARAGAYTNGARBMGNACNTTGGCGARATHTAARCCNTTNTAYGRCAYTN 720
Qy 241 HisAlaTyrValArgAlaLysLeuMetAsnAlaTyrProSerTyrIleSerProIleGly 260
Db 721 CAYGCNATYAGTNGMNMGNAAATYNTATGAYACNATYACCMNSNTAYATHTMSNCCNACGN 780
Qy 261 CysLeuProAlaHisLeuLeuGlyAspMetTrpGlyArgPheTrpHisAsnLeuTyrSer 280
Db 781 TGYTNTCCNGCNCAYTNTYNTNGNGAYATGTGGGNGMNTTYTGACNAAAYTNTYCCN 840
Qy 281 LeuThrValProPheGlyLysProAsnIleAspValThrAspAlaMetValAspGln 300
Db 841 YTNACNGTNCNTTYGNCARARARARARARARARARARARARARARARARARARARARAR 900

```

```

QY 301 AlaTrpAspAlaGlnArgIlePheLeuGluAlaGluLysPhePheValSerValGlyLeu 320
    |||||..... |||||..... |||||.....
Db 901 GGNNGGAGGCGNARGNATHTTTCAGANGCAGARARATTTTTCGMSNGNGAYTN 960
QY 321 ProAsnMetThrGlnGlyPheTrpGluAsnSerMetLeuThrAspProGlyAsnValGln 340
    |||||..... |||||..... |||||.....
Db 961 CCNCAATGACNARCGNTTGTGGCNAAYWSNATGTATNCNARCGNCGNAGYGNMG 1020
QY 341 LysAlaValCysHisProThrAlaIlePheValGlyLysGlyAspPheArgIleLeuMet 360
    |||||..... |||||..... |||||.....
Db 1021 AAKGTMTNTGTGACACCNACNGCNTGGGAYTTGNGCAYGNGAYTTTGMATTHAARANG 1080
QY 361 CysThrLysValThrMetAspPheLeuThrAlaHisGlnMetGlyHisIleGln 380
    |||||..... |||||..... |||||.....
Db 1081 TGACNARARCTNACNATGGGAYATTTTTCAGCAGCAGCAGCAGCAGCAGCAGCAG 1140
QY 381 TyrAspMetAlaTyrAlaAlaGlnProPheLeuLeuArgAsnGlyValAsnGlyPhe 400
    |||||..... |||||..... |||||.....
Db 1141 TAGGATATGCGCTAAGCNCARCCNTTYYTWTNMGNAAYGNCNNAAYGARGCNTTY 1200
QY 401 HisGluAlaValGlyIleMetSerLeuSerAlaAlaThrProLysHisLeuLysSer 420
    |||||..... |||||..... |||||.....
Db 1201 CAGGAGCNGCTNGCNGARARATHTGWSNTYTNWSNCGCNCNACNCAARCAAYTNARWGN 1260
QY 421 IleGlyLeuLeuSerProAspPheGlnGluAspAsnGluThrGluIleAsnPheLeuLeu 440
    |||||..... |||||..... |||||.....
Db 1261 ATGAGYNTYNTNCNMSNGAYTTTCARAGARAGATWSNGARACNARARATHTAATTTTNTYN 1320
QY 441 LysGluAlaLeuThrIleValGlyThrLeuProPheThrTyrMetLeuGluLysTrpArg 460
    |||||..... |||||..... |||||.....
Db 1321 AARCAAGCNYTNACNATHTGNGNACNYTNCNTYACNTAATGTCARARARATGCGMN 1380
QY 461 TrpMetValPheLysGlyIleLeuProLysAspGlnTrpMetLysTrpTrpGluMet 480
    |||||..... |||||..... |||||.....
Db 1381 TGGATGCTTTCGNGCNGARARATHTCCNARARARATGATGATGATGATGATGATGATG 1440
QY 481 LysArgGluIleValGlyValAlaGluProValProHisAspGluThrTyrCysAspPro 500
    |||||..... |||||..... |||||.....
Db 1441 AARMGARARATHTGNGNCTNGNARCCNYTNCNCAAYGARGARCNATYTCGAYCCN 1500
QY 501 AlaSerLeuPheHisValSerAsnAspTyrSerPheIleArgTyrThrArgThrLeu 520
    |||||..... |||||..... |||||.....
Db 1501 GCMMSYTTTTCATGATMSNAAYGATYTWSTTTCATHTMGTATTAAYACNMGNACNATH 1560
QY 521 TyrGlnPheGlnPheGlnGluAlaLeuCysGlnAlaAlaLysHisGluGlyProLeuHis 540
    |||||..... |||||..... |||||.....
Db 1561 TAYCARTTYCARTTYCARGARCCNTNTGTCARGCNGCNAARATAYAAAGMSNTTNCAY 1620
QY 541 LysCysAspIleSerAsnSerThrGluAlaGlyGlnLysLeuPheAsnMetLeuArgLeu 560
    |||||..... |||||..... |||||.....
Db 1621 AARTGAYATHTMSNAAYWSNACNGCNGCARAARATYTNARATGYTNMSNTYN 1680
QY 561 GlyLysSerGluProThrPheLeuAlaGluAsnValAlaGlyAlaLysAsnMetAsn 580
    |||||..... |||||..... |||||.....
Db 1681 GGNAAAYWSNGARCCNTGACNARCCNTNGARAAAYGNTGNGCNGNMGNAAYATGAY 1740
QY 581 ValArgProLeuLeuAsnTyrPheGluProLeuPheThrTrpLeuLysAspGlnAsnLys 600
    |||||..... |||||..... |||||.....
Db 1741 GTNARCCNYTNTAAAYTATYTCARCCNTNTTTCAYTGGYTAAARGARCARAAVMGN 1800
QY 601 AsnSerPheValGlyTyrSerThrAspTrpSerProTyrAlaAspGlnSerIleLysVal 620
    |||||..... |||||..... |||||.....
Db 1801 AAYWSNTTTCGNTGGAAGAAATGAGWSNCCNTAYGCGNAYCARMSNATTHAARCTN 1860
QY 621 ArgIleSerLeuLysSerAlaLeuGlyAspLysAlaTyrGluTrpAsnAspAsnGluMet 640
    |||||..... |||||..... |||||.....
Db 1861 MGNATHWSNTYTAARWSNCGNTTNGCNCNAAATGACNTAYGARTGACNAAATGARTG 1920
QY 641 TyrLeuPheArgSerSerValAlaTyrAlaMetArgGlnTyrPheLeuLysValLysAsn 660
    |||||..... |||||..... |||||.....
Db 1921 TTYTNTTTCGMSNMSNCGNTGNTAYGNCNATGMAARATYATYTWSNATHTAARAAAY 1980
QY 661 GlnMetIleLeuPheGlyGluGluAspValArgValAlaAsnLeuLysProArgIleSer 680

```

```

Db 1981 CARACNGTNCCTTYYTNARGARGAGATNMGNGTNSGAYTTNARCCNMGNGTNSN 2040
QY 681 PheAsnPhePheValThrAlaProLysAsnValSerAspIleIleProArgThrGluVal 700
    |||||..... |||||..... |||||.....
Db 2041 TTYTATYTTTTCGTHACNMSNCCNCAARAYGNTMSNGAYGNTAATHTCCNMGMSNGARCTN 2100
QY 701 GluLysAlaIleArgMetSerArgIleAsnAspAlaPheArgLeuAsnAspAsn 720
    |||||..... |||||..... |||||.....
Db 2101 GARGAYGNCNATHMGNATGMSNMGNGNMGNTHTAAYGAYGNTTTCGNTYTAAGATYAY 2160
QY 721 SerLeuGluPheLeuGlyIleGlnProThrLeuGlyProProAsnGlnProProValSer 740
    |||||..... |||||..... |||||.....
Db 2161 WSNYTNAGATYYTNGNATHTCAACNANTNGARCCNCCNTAYCARCCNCCNCTNACN 2220
QY 741 IleTrpLeuIleValPheGlyValAlaMetGlyValIleValGlyIleValIleVal 760
    |||||..... |||||..... |||||.....
Db 2221 AHTGTGYNATHATHTTTCGNGTNGTNTATGCGCTTNGTNGTNGTNGTNGTNGTNGTNG 2280
QY 761 IlePheThrGlyIleArgAspArgLysLysLysAsnLysAlaArgSerGlyLysAsnPro 780
    |||||..... |||||..... |||||.....
Db 2281 AHTGTNACNGNATHAARGNMGNAARAARAARARACNAAARMGARGARARAYCCN 2340
QY 781 TyrAlaSerIleAspIleSerLysGlyLysAsnAsnProGlyPheGlnAsnThrAspAsp 800
    |||||..... |||||..... |||||.....
Db 2341 TAYGAYWSNATGATATHTGNAARAGNARWSNAAAYCGNGNTTTCARAAAYWSNGAYAY 2400
QY 801 ValGlnThrSerPhe 805
    |||||..... |||||.....
Db 2401 GCMNARACNMSNTTY 2415

```

Search completed: March 7, 2003, 11:03:34  
Job time : 285 secs

